

UNIVERSITÉ DU QUÉBEC À MONTRÉAL

INTELLIGENCE ARTIFICIELLE ET ROBOTIQUE BIO-INSPIRÉE :  
MODÉLISATION DE FONCTIONS D'APPRENTISSAGE  
PAR RÉSEAUX DE NEURONES À IMPULSIONS

THÈSE  
PRÉSENTÉE  
COMME EXIGENCE PARTIELLE  
DU DOCTORAT EN INFORMATIQUE COGNITIVE

PAR  
ANDRÉ CYR

SEPTEMBRE 2016

UNIVERSITÉ DU QUÉBEC À MONTRÉAL  
Service des bibliothèques

Avertissement

La diffusion de cette thèse se fait dans le respect des droits de son auteur, qui a signé le formulaire *Autorisation de reproduire et de diffuser un travail de recherche de cycles supérieurs* (SDU-522 – Rév.07-2011). Cette autorisation stipule que «conformément à l'article 11 du Règlement no 8 des études de cycles supérieurs, [l'auteur] concède à l'Université du Québec à Montréal une licence non exclusive d'utilisation et de publication de la totalité ou d'une partie importante de [son] travail de recherche pour des fins pédagogiques et non commerciales. Plus précisément, [l'auteur] autorise l'Université du Québec à Montréal à reproduire, diffuser, prêter, distribuer ou vendre des copies de [son] travail de recherche à des fins non commerciales sur quelque support que ce soit, y compris l'Internet. Cette licence et cette autorisation n'entraînent pas une renonciation de [la] part [de l'auteur] à [ses] droits moraux ni à [ses] droits de propriété intellectuelle. Sauf entente contraire, [l'auteur] conserve la liberté de diffuser et de commercialiser ou non ce travail dont [il] possède un exemplaire.»

UNIVERSITY OF QUEBEC AT MONTREAL

ARTIFICIAL INTELLIGENCE AND BIO-INSPIRED ROBOTICS:  
LEARNING FUNCTIONS MODELED WITH  
SPIKING NEURAL NETWORKS

THESIS  
SUBMITTED  
IN PARTIAL FULFILMENT OF THE REQUIREMENTS  
FOR THE DEGREE OF  
DOCTOR IN COGNITIVE INFORMATICS

BY  
ANDRÉ CYR

SEPTEMBER 2016





## REMERCIEMENTS

À mes amis qui m'ont toujours encouragé dans mon étrange deuxième vie.

À Jean Boucher, professeur au département de kinanthropologie de l'UQAM, qui a vraiment été le premier à me donner (entre autres) la chance de côtoyer le monde de la recherche.

À Réjean Dubuc, professeur au département de kinanthropologie de l'UQAM et de neurosciences à l'Université de Montréal, qui m'a appris les rudiments du métier de chercheur lors de la maîtrise en neurosciences. J'en garde un souvenir impérissable, mais aussi un grand regret de l'avoir laissé tomber en ne terminant pas mon mémoire. J'espère qu'en ce jour, j'aurai su rendre honneur à son dévouement...

À Marie-Ève Clément, professeur au département de psychologie à l'UQO et mère de mes deux enfants. Véritable source d'ambition et modèle de travail acharné, c'est elle qui a le mérite de m'avoir remis sur le chemin des études universitaires.

À Pierre Poirier, professeur au département de philosophie et informatique cognitive à l'UQAM, qui a joué un rôle déterminant comme codirecteur dans l'initiative et l'orientation de cette thèse. Avec ses suggestions critiques du domaine de l'IA et par nos nombreuses conversations, il a ainsi fortement influencé le sujet principal de ce doctorat et établi les fondements de mes futures recherches. Merci Pierre !

À Mounir Boukadoum, professeur au département d'informatique de l'UQAM, principal codirecteur m'ayant soutenu dans le grand droit de ma thèse. Mounir m'a accompagné avec rigueur dans les articles qui ont été produits. Il a su respecter mes

idées et les étapes de mon cheminement, si atypique fût-il. Aussi, il m'a conduit vers le chemin de l'autonomie comme chercheur. Merci Mounir !

À mes deux codirecteurs, pour leur aide dans tous mes écrits. J'espère pouvoir maintenant dire que mon style d'écriture aura à tout le moins, suivi une courbe d'apprentissage positive durant cet intervalle de temps doctoral !

À Frédéric Thériault, engagé comme soutien technique au départ du projet SIMCOG, ce véritable génie informatique a toujours eu la capacité de transformer mes pensées en outils logiciels performants. Sans lui, cette thèse telle qu'elle l'est n'aurait jamais eu lieu. À ma grande joie, il est devenu un collègue de recherche et un partenaire dans l'élaboration de la suite logicielle SIMCOG. Frédéric est surtout devenu un ami avec le temps. Si je termine ce doctorat aujourd'hui, c'est un peu à cause de sa persévérance à m'encourager à le faire. Je vous souhaite vraiment une longue carrière dans le monde de la recherche...

À l'UQAM, en tant qu'institution libre et démocratique, la souplesse de ses règles administratives et son côté avant-gardiste auront permis à un étudiant différent de compléter un diplôme d'études supérieures.

À mes enfants, qui auront certainement le souvenir d'un père encore sur « l'ordi », mais qui je l'espère sauront que les études universitaires représentent bien plus qu'un diplôme conduisant à un travail...

À ma femme Tracy, qui malgré une certaine incompréhension qui demeure, me supporte et subit ma folie de tous les instants. Sans elle, je n'aurai jamais eu l'occasion de faire ce doctorat, suppléant à tous les besoins de la famille, alors qu'égoïstement je prenais du temps pour ma passion. Je ne serai jamais assez redevable, merci Tracy. Je t'aime mon amour !

## DÉDICACE

À mon père, qui j'en suis sûr a toujours été fier de moi. Malheureusement, le destin aura choisi de ne pas lui laisser assez de temps, pour assister à l'aboutissement de ce projet. Tu as souvent été seul à croire en tes rêves et c'est bien ta perspicacité qui les a conduits à terme. Maintenant, je te comprends...



**“The measure of intelligence is the ability to change.”**

**Albert Einstein**



## AVANT-PROPOS

Le thème original de ce doctorat a forcément évolué depuis le début de mon inscription, il y a un peu plus d'une douzaine d'années maintenant. Lors de ma présentation prédoctorale, le sujet central portait sur un lien possible entre les multiples facettes de la temporalité et son implication dans le phénomène de l'intelligence. Cette hypothèse demeure toujours en filigrane de l'actuelle thématique proposée, mais ne sera plus développée de façon concrète dans ce cadre doctoral.

Aussi, le format déposé est maintenant devenu celui d'une thèse par articles. En effet, trois articles déjà publiés sont maintenant liés par un seul thème dans ce doctorat (voir annexe B, C et D). En conséquence, la structure du document normalement attendue change légèrement. Les chapitres de méthodologie et d'analyse des résultats se retrouvent plutôt à l'intérieur de chacun des articles. Dans cette même perspective, l'introduction offre une mise en contexte théorique détaillée, toutefois organisée en une séquence logique. Il s'agit d'un complément à la revue de la littérature de chacun des articles, permettant de mieux les intégrer dans la thématique globale de la thèse. Cette introduction représente le condensé d'une réflexion longitudinale nécessaire, ayant permis d'aboutir au sujet actuel. Elle reflète aussi l'ampleur et la convergence de connaissances de plusieurs domaines différents, ayant significativement contribué au choix final du thème de ce doctorat en informatique cognitive. Incidemment, par mon bagage académique antérieur en médecine et en neurosciences, le modèle computationnel d'apprentissage proposé implique un accent sur le plan biologique.

De façon parallèle, un outil logiciel unique (SIMCOG: IA-Futur) a été développé pour faciliter l'élaboration des articles de la thèse. En effet, aucun logiciel ne rassemble actuellement la capacité d'éditer rapidement des architectures de réseaux de neurones à impulsion qui sont complexes, asymétriques et hétérogènes en propriétés neuronales



et un transfert rapide de ce contrôleur vers un monde virtuel 3D dédié au bio-inspiré ou vers des plateformes physiques de robots. Cette suite logicielle est détaillée dans l'annexe A, tout comme l'ajout de deux autres articles scientifiques qui ont été produits (annexe F et G). Bien que ces lectures supplémentaires soient facultatives en regard du sujet principal, elles offrent une meilleure compréhension de l'outil qui est utilisé dans les articles de la thèse.

Également, un quatrième article scientifique a été produit et ajouté (voir annexe E). Bien que non essentiel, cet article complète, résume et oriente la thèse vers une thématique d'apprentissage de plus haut niveau pour des recherches futures.

Enfin, ce doctorat représente l'aboutissement d'un travail soutenu et passionné. Par cette contribution du domaine en informatique cognitive, j'espère sincèrement avoir pu atteindre l'objectif doctoral d'acquérir la démarche scientifique conduisant au rôle de chercheur autonome. Chers membres du jury, je vous souhaite une agréable lecture...



## TABLE DES MATIÈRES

DÉDICACE .....	vii
AVANT-PROPOS .....	xi
LISTE DES FIGURES .....	xvii
LISTE DES ABRÉVIATIONS .....	xix
RÉSUMÉ .....	xxi
ABSTRACT .....	xxiii
CHAPITRE I	
INTRODUCTION .....	1
1.1 Contexte général .....	1
1.2 Problématique, hypothèse et méthodologie .....	2
1.3 Plan de la thèse.....	5
1.4 Intelligence artificielle .....	6
1.5 Robotique bio-inspirée.....	11
1.6 Agents autonomes complets virtuels et physiques.....	12
1.7 Temporalité.....	16
1.8 Réseaux de neurones artificiels.....	19
1.9 Réseaux neuronaux artificiels dynamiques classiques.....	25
1.10 Réseaux neuronaux artificiels à impulsions ou pulsés .....	26
1.11 Cognition et apprentissage en robotique .....	36
CHAPITRE II	
HABITUATION .....	41
2.1 Volet neuroscience .....	41
2.2 Volet IA .....	48
2.3 Thématiques connexes à explorer .....	51
CHAPITRE III	
CONDITIONNEMENT CLASSIQUE .....	57
3.1 Volet psychologie .....	57
3.2 Volet neuroscience .....	58
3.3 Volet IA .....	66
CHAPITRE IV	
CONDITIONNEMENT OPÉRANT .....	69
4.1 Volet psychologie .....	69

4.2	Volet neuroscience .....	72
4.3	Volet IA.....	73
CHAPITRE V		
DISCUSSION .....		77
5.1	Robotique bio-inspirée et apprentissage.....	77
5.2	RNAI .....	79
5.3	Habituatation .....	83
5.4	Conditionnement classique .....	84
5.5	Conditionnement opérant .....	86
CHAPITRE VI		
CONCLUSION ET PERSPECTIVES .....		89
6.1	Directions futures .....	91
6.1.1	Évolution verticale des champs d'expertise déjà développés.....	91
6.1.2	Évolution horizontale des champs de la cognition en IA .....	92
6.1.3	Standards de tests .....	93
6.2	Autre article paru et complémentaire de la thèse .....	94
6.2.1	La sélection d'action .....	94
6.3	Articles en préparation .....	97
6.3.1	Le façonnement comportemental .....	97
ANNEXE A		
SIMCOG.....		99
A.1	NeuroCode .....	102
A.2	NeuroSim .....	109
A.3	NeuroData .....	116
A.4	NeuroSync.....	119
A.5	NeuroGene .....	120
ANNEXE B		
ARTICLE HABITUATION.....		125
ANNEXE C		
ARTICLE CONDITIONNEMENT CLASSIQUE .....		143
ANNEXE D		
ARTICLE CONDITIONNEMENT OPÉRANT .....		161
ANNEXE E		
ARTICLE SÉLECTION D'ACTIONS .....		175

ANNEXE F	
ARTICLE SIMCOG .....	187
ANNEXE G	
CHAPITRE DE LIVRE SUR LES MONDES VIRTUELS.....	205
BIBLIOGRAPHIE.....	227



## LISTE DES FIGURES

Figure	Page
1.1 Concepts dichotomiques principaux en intelligence artificielle	10
1.2 Le neurone formel	21
1.3 Exemple d'une architecture de RNA	23
1.4 Exemples de motifs de sorties avec les RNAI	30
1.5 Diagramme de divers types d'apprentissage	39
2.1 Le phénomène de l'habituation	41
3.1 STDP	61
4.1 Types de procédures dans le conditionnement opérant	71
4.2 Exemple des principales composantes expérimentales du CO	71
4.3 Motif cellulaire singulier capable de simuler un CO	74
6.1 Trois différents types d'apprentissage expliqués au niveau cellulaire	90
A.1 Schématisation de la structure de la suite logicielle SIMCOG	100
A.2 Algorithme décrivant les étapes fonctionnelles du logiciel SIMCOG	101
A.3 Exemple d'une saisie d'écran typique dans NeuroCode	104
A.4 Principales composantes de la variance d'un potentiel de membrane neuronal	107
A.5 Figure démontrant une fonction mathématique de type alpha	107
A.6 Exemple d'une saisie d'écran typique d'un RNAI réalisée avec NeuroCode	108
A.7 Exemple d'une saisie d'écran en vue panoramique réalisée avec NeuroCode	109
A.8 Schéma d'une séquence logique utilisée par un usager dans NeuroSim	111
A.9 Exemple d'une saisie d'écran dans NeuroSim	113
A.10 Exemple dans NeuroSim : exploration d'un labyrinthe par quatre robots	115

A.11 Exemple d'une saisie d'écran dans NeuroSim en vue orbitale	116
A.12 Exemple d'une saisie d'écran typique dans NeuroData	117
A.13 Exemple dans NeuroData, d'un comparatif de robots de la même simulation	118
A.14 Exemple dans NeuroData, d'une séquence logique temporelle pour un robot	119

## LISTE DES ABRÉVIATIONS

AG	Algorithmes génétiques
AMPA	Acide $\alpha$ -Amino-3-hydroxy-5-méthyl-4-isoxazolépropionique
IA	Intelligence artificielle
IAG	Intelligence artificielle générale
IIS	Intervalle interstimulus
CA1	Aire 1 de l'hippocampe (du latin: regio I cornus ammonis)
Ca <sup>++</sup>	Calcium
CC	Conditionnement classique
CO	Conditionnement opérant
CPU	Central processing unit
CR	Conditioned response
CS	Conditioned stimulus
IIB	Interval inter-bloc
IIS	Interval inter-stimulus
K <sup>+</sup>	Potassium
LIF	Leaky integrated-and-fire
LTD	Long-term depression
LTP	Long-term potentiation
Mg <sup>++</sup>	Magnésium
Ms	Milliseconde
MV	Millivolt
NMDA	N-Méthyl-D-Aspartate
NA	Nanoangström
NS	Neutral stimulus
PA	Potentiel d'action
PM	Potentiel de membrane
PPS	Potentiel postsynaptique
PPSE	Potentiel postsynaptique excitateur
PPSI	Potentiel postsynaptique inhibiteur
PR	Potentiel de repos
PRA	Période réfractaire absolue
RAM	Random-access memory
RNA	Réseaux de neurones artificiels
RNAI	Réseaux de neurones artificiels à impulsions
SRM	Spike response model
STDP	Spike-timing dependent plasticity
UR	Unconditioned response
US	Unconditioned stimulus





## RÉSUMÉ

Cette thèse a comme objectif de permettre une avancée originale dans le domaine de l'informatique cognitive, plus précisément en robotique bio-inspirée. *L'hypothèse défendue est qu'il est possible d'intégrer différentes fonctions d'apprentissage, élaborées et incarnées pour des robots virtuels et physiques, à un même paradigme de réseaux de neurones à impulsions agissant comme cerveaux-contrôleurs.*

La conception de règles d'apprentissage et la validation de l'hypothèse de recherche reposent sur la simulation de mécanismes cellulaires à base de plasticité synaptique et sur la reproduction de comportements adaptatifs des robots. Cette thèse par articles cible trois types d'apprentissage de complexité incrémentale : l'habituation comme forme d'apprentissage non associatif et les conditionnements classiques et opérants comme formes d'apprentissage associatif. L'analyse détaillée, de la synapse au comportement, est validée par des études expérimentales provenant d'invertébrés tels que le ver nématode *Caenorhabditis elegans*. Pour chacune de ces règles, un algorithme novateur a été proposé, conduisant à la publication d'un article scientifique.

Ces règles d'apprentissage ont été modélisées en développant certains paramètres temporels et des circuits neuronaux précis. Incidemment, la granularité du temps des réseaux de neurones à impulsions (RNAI) est établie au niveau du simple potentiel d'action plutôt qu'au niveau du taux moyen de décharge par unité de temps, comme c'est le cas pour les réseaux de neurones artificiels traditionnels. Cette propriété des RNAI s'est avérée être un atout suffisant pour préférer leur utilisation pour des robots évoluant dans le monde réel.

L'élaboration du modèle computationnel d'apprentissage pour des robots a requis de tester d'abord les hypothèses sur des simulations virtuelles. Puisqu'aucun simulateur n'avait les capacités suffisantes pour tester notre hypothèse, soit d'intégrer des RNAI, des structures de robots, et des interfaces pour l'exportation des RNAI vers des plateformes physiques et des environnements virtuels 3D suffisamment complexes, il a été nécessaire de développer, en parallèle de la thèse, un logiciel novateur (SIMCOG), permettant une étude analytique par le suivi dynamique des variables, des synapses de RNAI jusqu'aux comportements d'un ou plusieurs robots virtuels ou physiques.

Finalement, outre l'intégration de plusieurs fonctions différentes d'apprentissage dans des RNAI, une autre des conclusions de ce travail suggère que des robots virtuels et physiques peuvent apprendre et s'adapter au niveau comportemental, de façon similaire aux agents naturels. Ces observations comportementales sont basées sur la simulation de mécanismes de plasticité synaptique modulés par des variables temporelles relatives aux stimuli physiques et aux activités cellulaires neuronales.

**Mots-clés :** Intelligence artificielle, Cognition, Simulateur, Robotique bio-inspirée, Réseaux de neurones artificiels à impulsions, Apprentissage, Habituation, Conditionnement classique, Conditionnement opérant, Plasticité synaptique



## ABSTRACT

The main goal of this PhD thesis is to validate a set of original ideas in the cognitive informatics, and more specifically in the field of bio-inspired robotics. The main hypothesis is that it is possible to integrate different learning rules, embedded and built for virtual and physical robots, in an artificial spiking neural networks (ASNN) paradigm.

The conception of the learning rules and the validation of the hypothesis rest on the simulation of the natural cellular mechanisms of synaptic plasticity while reproducing adaptive behaviors at the level of the complete cognitive agents. This three-article thesis targets three types of learning function that are incremental in complexity: habituation as a form of non-associative learning, and classical and operant conditioning as forms of associative learning.

Detailed analyses, from synapses to behaviors, are based on invertebrate neural experimental models such as the worm *Caenorhabditis elegans*. For each learning rule, a novel algorithm and a scientific article was produced.

The learning functions were modelled by developing specific temporal parameters and neural circuits. Indeed, ASNN computes with a single spike resolution, instead of the rate-coding methods used by traditional artificial neural networks. This temporal property was enough to prefer these neural models for our robot's brain-controllers.

However, the computational ASNN model and the connected learning rules elaborated for robots required virtual simulations to test the research hypothesis. Since no simulators could easily link ASNN, robots structures, as well as transfer interfaces for virtual and physical robots and a virtual 3D world complex enough to test various learning scenarios, we developed a novel simulation package (SIMCOG), in parallel to the presented thesis. This software allows exhaustive analysis, from the dynamics of synapses to the behaviors of one or many virtual and physical robots.

Finally, beyond the integration of different learning rules in an ASNN, one of the conclusions drawn from this work is that robots can learn and show adaptive behaviors based on cellular mechanisms similar to those found in natural cognitive agents.

**Keywords:** Artificial intelligence, Cognition, Simulator, Bio-inspired robotics, Spiking neural networks, Learning, Habituation, Classical conditioning, Operant conditioning, Synaptic plasticity



## CHAPITRE I

### INTRODUCTION

#### 1.1 Contexte général

L'intelligence est un thème largement étudié dans plusieurs disciplines scientifiques. Malgré cet engouement pour le phénomène, la terminologie exacte ne fait aucunement l'unanimité et reflète probablement les différents angles abordés par les multiples experts. En sciences cognitives (Matlin, 2005), l'intelligence représente la conséquence de l'ensemble des processus mentaux de haut niveau ainsi que leurs interrelations. Pour ne nommer que ceux-ci, la mémoire, le raisonnement, la connaissance, l'apprentissage et le langage sont des thèmes communément explorés dans ce domaine.

Partageant une partie du spectre des sciences cognitives, l'intelligence artificielle (IA) comme discipline (Luger, 2002; Russell et Norvig, 2010), aspire aussi à comprendre le phénomène, mais, en tentant plutôt d'en reproduire certaines caractéristiques ciblées à travers des artéfacts non biologiques. Ces IA sont développées par différentes approches, dont celles de l'inspiration biologique. Un des buts de l'IA est de dépasser l'étude théorique du phénomène par la conception d'agents cognitifs pouvant, par exemple, prendre l'aspect d'un robot. Un objectif partagé par l'IA et la robotique est de réaliser un jour, un artéfact physique autonome exhibant les caractéristiques essentielles de la cognition humaine.

Dans cette perspective, certains chercheurs (Pfeifer et Scheier, 1999) définissent l'intelligence comme étant le résultat observé par une personne externe selon des critères précis, d'un agent incarné percevant et agissant dans son monde par des comportements adaptatifs. Il devient ainsi possible de comprendre le phénomène général, autant dans le monde naturel que dans le monde artificiel.



Une autre avenue traditionnelle d'exploration de l'intelligence est celle des neurosciences, où le neurone biologique est souligné comme élément central dans le phénomène. Incidemment, l'étude du cerveau, du niveau moléculaire et génétique au niveau comportemental, passe par un éventail de techniques d'investigation aidant à la compréhension globale de l'intelligence (Kandel *et al.* 2012).

C'est dans ce contexte d'intégration, au carrefour des connaissances en science cognitive, IA, neuroscience et robotique, que s'inscrit cette thèse en informatique cognitive. L'apprentissage en neuro-robotique est le sujet et la contribution principale. Un modèle computationnel d'inspiration biologique est présenté, intégrant trois fonctions d'apprentissages primaires, développées pour des robots avec une perspective globale d'étude de l'intelligence en tant que phénomène général. Ce modèle unifié d'apprentissage aspire à la réalisation d'agents cognitifs artificiels plus complexes, qui rivaliseront peut-être un jour avec ceux de l'intelligence naturelle.

## 1.2 Problématique, hypothèse et méthodologie

Cette thèse en informatique cognitive est articulée spécifiquement dans le domaine de la neuro-robotique par une approche biologiquement inspirée. L'hypothèse H1 représente le thème central et ce qui a été démontré alors que l'hypothèse H2 représente une intuition des processus sous-jacents, qui pourrait être explorée ultérieurement :

***H1 : Il est possible de reproduire différents types d'apprentissage biologique dans un même paradigme de réseau de neurones artificiels à impulsions (RNAI) agissant comme cerveau-contrôleurs de robots virtuels et physiques.***

***H2 : La cohérence du modèle computationnel implémenté pour des agents complets situés et incarnés émerge des relations temporelles qui existent entre les stimuli physiques, les réponses neuronales, les mécanismes de plasticité synaptique des règles d'apprentissages et l'adaptation comportementale observable modifiée par l'expérience.***

L'apprentissage est évalué au niveau comportemental des robots avec des données empiriques comparables à celles des modèles animaux. L'application pratique de cette étude concerne le développement de capacités adaptatives pour des robots, résolues par l'implémentation de fonctions d'apprentissage reposant sur des mécanismes cellulaires similaires à ceux observés avec les agents biologiques.

Afin de parvenir aux résultats, des propriétés critiques structurelles et fonctionnelles soutenant l'apprentissage ont été ciblées, en s'inspirant de modèles animaux neuronaux simples, mais de complexité incrémentale. Les organismes biologiques tels que le ver nématode *Caenorhabditis elegans*, la limace de mer *Aplysia californica*, la mouche *Drosophila melanogaster*, ainsi que l'abeille *Apis mellifera* représentent des choix qui font consensus dans la communauté scientifique et le domaine de l'apprentissage (Menzel et Benjamin, 2013). Ces modèles animaux ont permis d'étudier à de multiples niveaux d'analyse, de la cellule au comportement, différents types d'apprentissage. L'approche méthodologique de cette thèse a été de reproduire certaines fonctions primaires d'apprentissage, dont l'habituation, le conditionnement classique (CC) et le conditionnement opérant (CO). Le niveau de biomimétisme visé pour la réalisation des différents types d'apprentissage a été celui des éléments cellulaires neuronaux, particulièrement la composante synaptique des RNAI. L'efficacité des algorithmes et des simulations a été continuellement ajustée avec la rétroaction des résultats obtenus au niveau comportemental des robots.

Pour élaborer les algorithmes, il a fallu sélectionner certaines variables des processus biologiques soutenant ces règles d'apprentissage pour ensuite les modéliser et les intégrer dans un RNAI. Des aspects novateurs ont été produits pour chacune de ces règles. Pour l'habituation, l'application de la fonction primaire a été étendue dans le domaine temporel. L'algorithme proposé intègre des réponses variables aux motifs entrants de stimuli asynchrones, aux différentes magnitudes des stimuli ainsi qu'aux effets de potentialisation entre les différents temps de sessions d'apprentissage. Pour le CC, un mécanisme interne associatif de type *spike-timing dependent plasticity* (STDP) a été proposé avec un algorithme permettant des réponses synaptiques modulées en

relation avec le pairage temporel pré-postsynaptique, mais aussi par l'historique des fréquences individuelles des potentiels d'action (PA) pré et post ainsi que la fréquence des paires pré-post. Pour le CO, cette thèse propose un mécanisme d'apprentissage basé sur un motif singulier, composé de quelques cellules, synapses et des fonctions d'apprentissage précédemment développées. Ce motif a été suffisant pour reproduire plusieurs scénarios canoniques associés avec le CO. L'ensemble de ces fonctions d'apprentissages a été implémenté et validé dans des agents robotisés complets, virtuels et physiques.

Une revue de la littérature des différents domaines a été effectuée tout au long des recherches en vue et de la rédaction de la thèse, afin d'acquérir l'expertise nécessaire conduisant à la réalisation de l'ensemble de ces trois fonctions d'apprentissages. Une des conclusions de cette recherche a été de constater que très peu d'études jumellent les RNAI, l'apprentissage et la robotique. Ce sujet doctoral a donc démarré à partir de résultats expérimentaux et des lacunes observées dans ces différents domaines.

Ainsi, le survol de ces thématiques a révélé l'absence d'outils permettant de résoudre adéquatement la problématique d'intégration de multiples fonctions d'apprentissage à des RNAI visant des applications de robots virtuels ou physiques. Le manque de logiciels facilitant l'élaboration rapide de RNAI variés, couplés à des simulateurs de robots, a ainsi justifié le besoin d'innover pour répondre sur mesure aux solutions anticipées. Ce qui a conduit au développement de SIMCOG, une suite logicielle intégrant un éditeur de RNAI (NeuroCode) ainsi qu'un éditeur et simulateur de mondes virtuels 3D (NeuroSim). NeuroCode permet aussi l'exportation des architectures de RNAI vers plusieurs plates-formes physiques par le biais de simples interfaces. SIMCOG possède aussi un outil de visualisation (NeuroData) des données en temps différé et en temps réel. Enfin, dans une optique de travail académique, des avatars permettent à plusieurs utilisateurs d'éditer des RNAI et des mondes virtuels en mode partagé à distance.

Dans cette thèse, SIMCOG a permis d'intégrer différents niveaux de conception d'IA bio-inspirées, facilitant la validation d'hypothèses reliées aux fonctions



d'apprentissages à travers l'incarnation de divers robots virtuels et physiques. Le résultat est que ces règles d'apprentissage peuvent être mimées pour des robots situés et incarnés, dans un paradigme de RNAI et résolues par des mécanismes de plasticité synaptique similaires à ceux du monde biologique.

Bien que SIMCOG se soit avéré un élément essentiel pour parvenir efficacement à la validation des hypothèses de recherche formulées ci-dessus, son développement ne constitue pas le cœur de cette thèse, mais une bonification facultative de la composante informatique doctorale.

### 1.3 Plan de la thèse

L'organisation de cette thèse s'établit comme suit : tout d'abord, l'introduction se poursuit par un survol nécessaire des domaines de la robotique bio-inspirée, de l'IA, des sciences cognitives et des neurosciences, contextualisant plus en profondeur l'aboutissement et la réflexion vers le sujet principal. La section suivante décrit les agents mobiles autonomes virtuels et incarnés comme étant les cibles appliquées, justifiant le développement d'un nouveau modèle intégrateur de fonctions d'apprentissage. Ensuite, une sous-section sur la temporalité explique le déclencheur et le ligand des différents concepts dans cette thèse. Enfin, ce premier chapitre se termine par une revue générale des RNA, précisant le choix des RNAI comme modèle pour l'implémentation des fonctions d'apprentissage.

Le deuxième chapitre concerne l'apprentissage en robotique. Le cœur de la thèse y est développé avec l'élaboration de règles d'apprentissage bio-inspirées pour des robots. Ce chapitre inclut l'habituation, le conditionnement classique (CC) et le conditionnement opérant (CO). Pour chacune de ces fonctions d'apprentissage, un article a été produit, que l'on retrouve en annexe. Finalement, cette thèse doctorale aura démontré que des robots virtuels et physiques peuvent simuler de l'habituation, du CC et du CO. La contribution au domaine informatique aura été de développer un modèle computationnel unificateur de ces algorithmes dans un paradigme de RNAI. La

contribution en sciences cognitives aura été de simuler ces mécanismes d'apprentissage naturels au niveau fonctionnel de la plasticité synaptique, en reproduisant des résultats au niveau comportemental, comparable à ceux retrouvés avec les agents cognitifs biologiques.

La discussion suit au chapitre suivant où l'intégration des différentes hypothèses et démonstration des résultats obtenus sont réunies dans une perspective critique, contextualisée par le domaine de la robotique bio-inspirée. Dans ce chapitre, un lien est soulevé entre les fonctions d'apprentissage, la temporalité, les stimuli physiques et l'adaptation comportementale des agents cognitifs.

Ensuite, un chapitre est consacré à la conclusion et aux futures directions de recherches, prélude à l'élaboration d'un laboratoire en neuro-robotique. Un article complémentaire sur la sélection d'action a été ajouté dans cette section, puisque produit lors du doctorat, et qui intègre toutes les fonctions d'apprentissage de la thèse.

Puisque cette thèse et les articles produits utilisent le logiciel SIMCOG, un dernier chapitre facultatif y est consacré, sans pour autant détailler exhaustivement toutes les fonctionnalités disponibles. Un bref tour de l'éditeur de RNAI (NeuroCode) ainsi que l'éditeur et simulateur de mondes virtuels (NeuroSim), permettra au lecteur de mieux cerner le projet global et la pertinence de développer un tel logiciel. Ce chapitre se termine avec l'insertion d'un article et d'un chapitre de livre électronique paru sur cette suite logicielle. Enfin, les sections références et annexes clôturent cette thèse doctorale par articles.

#### 1.4 Intelligence artificielle

Depuis longtemps, l'homme rêve de fabriquer une machine pensante. L'arrivée des premiers ordinateurs a dramatiquement accentué cet intérêt et permis de concrétiser le projet autour d'une nouvelle discipline scientifique vers la fin des années 1950, l'IA. Pour les fondateurs de l'IA, John McCarthy, Allen Newell, Nathaniel Rochester,

Marvin Lee Minsky, Herbert Simon, Alan Turing et Claude Shannon, une machine serait perçue comme intelligente si elle reproduisait le comportement d'un humain.

Par la suite, certains chercheurs du domaine ont exploré l'IA conceptualisée du haut vers le bas, à l'aide de représentations internes et centrales, afin que ces machines puissent percevoir et agir par raisonnement sur le monde. L'explosion combinatoire a vite limité la capacité computationnelle à abstraire symboliquement les objets du monde et à les définir explicitement. Ce système symbolique caractérise bien le courant de pensée cognitiviste. Pour d'autres, l'IA pouvait se réaliser par des unités simples interconnectées, émulant les principes de base auto-organiseurs d'une cellule neuronale naturelle; ce qui définit le connexionnisme. Ces deux approches fondamentales et complémentaires reposent sur des hypothèses de travail différentes.

Dans les années 1970, plusieurs éléments et objectifs de l'IA ont été remis en question. Entre autres, que les caractéristiques de la vie et de la conscience ne pourraient être reproduites dans des structures non biologiques ! Une décennie plus tard, avec Rodney Brooks, l'approche de l'IA a opéré une transition majeure, plus pragmatique et sans représentation symboliquement définie, en proposant des modèles d'IA basés sur une coordination sensori-motrice simple, mais incrémentale en complexité (Brooks, 1986; 1991a; 1991b). Cette approche visait le niveau cognitif atteint par celui des insectes (Brooks, 1999).

La qualité de ce paradigme conceptualisant l'IA du bas vers le haut a été de recadrer le champ d'expertise vers des objectifs plus modestes et réalistes. L'architecture des contrôleurs, décomposée en multiples couches classifiées et hiérarchisées par fonctions et comportements, était née. La robotique par comportement hiérarchisé possède un certain degré de similarité avec le monde naturel ; le développement phylogénique et ontogénique est orchestré par le moteur évolutif de la sélection naturelle et les principes fondamentaux de la génétique. Cette approche est cependant encore limitée par les assises mêmes définissant les comportements à reproduire, parce que la complexité et la gestion de priorités des comportements par critères sont aussi des éléments qui atteignent rapidement une explosion combinatoire (Crevier, 1997).

Parallèlement, l'IA comme discipline est aussi fondée sur deux axes de recherche, soit une science visant la construction d'artéfacts qui exhibent des comportements intelligents et qui sont généralement conçus pour accomplir des tâches précises (IA classique ou faible) ou, soit vers l'émulation d'agents complets cognitivement comparables aux humains, caractérisant l'intelligence artificielle générale (IAG) ou l'IA forte<sup>1</sup>. Dans le premier cas, par exemple, un programme de reconnaissance de caractères est perçu comme une forme simple d'IA, où l'appariement d'un symbole avec un autre incomplet ou brouillé, accompagné d'un processus d'optimisation par apprentissage, représente le mimétisme d'un processus cognitif spécifique, soit la catégorisation. Autre exemple, le robot-rover envoyé sur Mars est qualifié d'IA faible, mais complexe, car à des degrés variables, son autonomie décisionnelle et ses capacités adaptatives miment certaines fonctions biologiques cognitives supérieures. Pour l'IAG, l'émulation vise une cognition plus large, en ciblant des caractéristiques de haut niveau conceptuel telles que le langage, les émotions et l'apprentissage sous toutes ses formes.

Enfin, l'IA est aussi scindée en deux groupes conceptuels. Certains scientifiques adhèrent au principe d'incarnation, et supportent comme prémisse une IA située dans le temps et dans un environnement, basée sur une structure physique. Il s'agit d'une IA ayant des capacités adaptatives cohérentes en regard de la triade corps-contrôleur-environnement (Beer, 1995; Chiel et Beer, 1997; Clark, 1997; Chrisley, 2003; Pfeifer, 2007, Ziemke, 2003b). Ce champ de la recherche caractérise bien la robotique bio-inspirée avec la vision d'une IA qui est incarnée. Dans le second groupe, les agents virtuels et programmes logiciels étendent la définition et remplissent partiellement certains critères d'une IA. Autour de ces différents paliers taxinomiques gravitent aussi des disciplines émergentes qui alimentent le domaine de l'IA, tel que celui de la vie artificielle et de la robotique développementale.

---

<sup>1</sup> Dans son sens initial (Searle, 1980), l'IA forte réfère à la thèse selon laquelle la manipulation de symboles, ou computation, serait une condition suffisante pour posséder de l'intelligence. Dans un sens commun et pour cette thèse, nous comprendrons l'IA forte comme référant à la possibilité de dupliquer dans un même système, l'ensemble des capacités générales de la cognition humaine avec des ordinateurs (Turing, 1950).

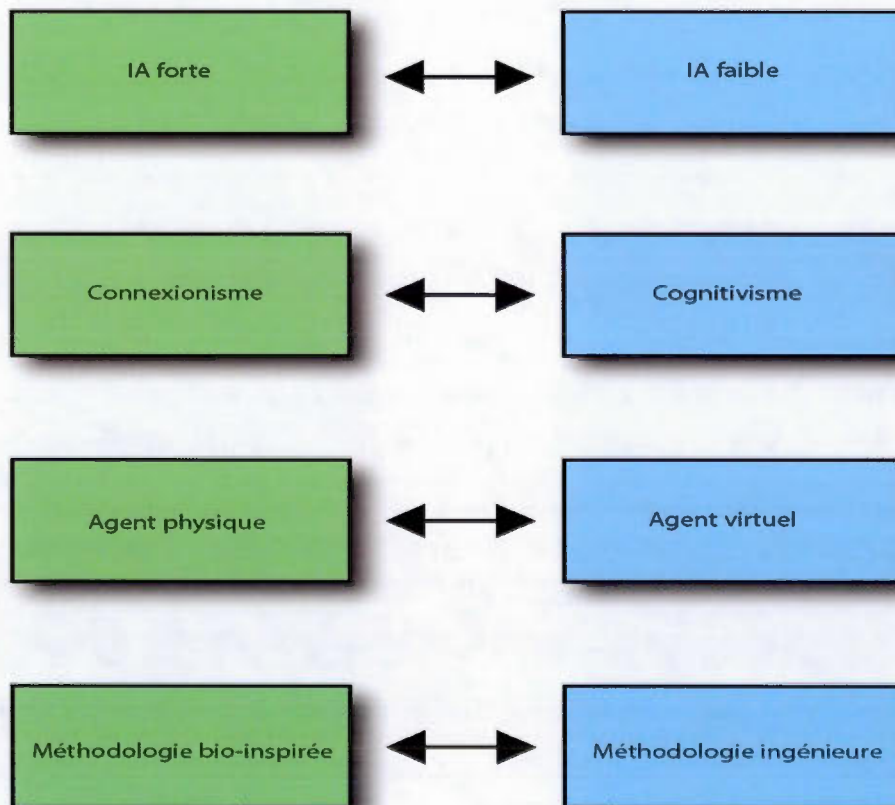
L'implémentation d'IA, et plus spécifiquement de contrôleurs, requiert une méthodologie de travail. À cet effet, les mathématiques en ingénierie, les algorithmes informatiques et les programmes symboliques formels servent traditionnellement de base conceptuelle. Aujourd'hui, plusieurs méthodes computationnelles sont également disponibles et complètent l'arsenal des outils. Une liste non exhaustive de celles-ci comprend notamment les algorithmes évolutionnaires, dont les algorithmes génétiques (AG), les réseaux bayésiens, la logique floue, les ondelettes et l'optimisation par essais particuliers. Somme toute, ces dernières méthodes sont basées sur des algorithmes et des mathématiques qui permettent d'inférer, d'approximer et d'optimiser des résultats à obtenir en fonctions de données entrantes et de critères de sélectivité des données sortantes.

Comme alternative méthodologique de développement de contrôleurs d'IA, la bio-inspiration offre aussi des algorithmes basés sur des programmes. Cette voie de recherche tend à maximiser le mimétisme du monde biologique sur de multiples niveaux conceptuels. Les RNAI peuvent ainsi être considérés dans cette catégorie, cumulant plusieurs caractéristiques majeures des systèmes neuronaux naturels, comme une communication par émission de PA et une plasticité synaptique reposant sur de multiples règles d'apprentissage. Aussi, les RNAI pourraient être conceptualisés par une phase développementale reposant sur des gènes artificiels et une modulation épigénétique secondaire à une évolution des environnements.

Dans cette perspective d'IA bio-inspirée, les contrôleurs de robots à base de RNAI (Wang *et al.* 2008) sont divisés en deux approches conceptuelles distinctes. Les neurosciences computationnelles étudient le cerveau avec l'aide de modèles mathématiques détaillés qui reproduisent les différentes activités neuronales; l'implémentation de modèles théoriques dans des robots physiques vérifie ensuite la plausibilité des hypothèses. À l'opposé, la neuro-robotique vise d'abord à réaliser des agents artificiels intelligents, mais qui utilise une modélisation fonctionnelle des neurones et des comportements.



En résumé (Figure 1.1), l'IA cherche à comprendre et réaliser des agents cognitifs artificiels à l'aide de théories diverses et de multiples méthodologies. Cette thèse soutient la théorie connexionniste avec les RNAI et le concept d'une IA générale, en mire avec l'éventuelle réalisation d'un agent complet, autonome et non asservi (blocs verts de la Figure 1.1). Dans ce doctorat, la méthodologie utilisée pour réaliser et étudier l'IA, incline préférentiellement vers l'implémentation de robots physiques, mais par souci d'efficacité et d'optimisation, assume au préalable l'élaboration partielle des modèles par des simulations virtuelles. Enfin, dans le développement de contrôleurs pour des robots, cette thèse favorise une stratégie d'inspiration biologique, spécifiquement pour ce qui est de la réalisation des fonctions d'apprentissages, à travers des algorithmes de gestion des synapses artificielles de RNAI.



**Figure 1.1** Concepts dichotomiques principaux en intelligence artificielle.

### 1.5 Robotique bio-inspirée

La robotique bio-inspirée est un champ émergeant de l'IA (Bekey, 2005; Holland, 2003; Floreano et Mattiussi, 2008; Liu et Sun, 2012; Lakhtakia et Martin-Palma, 2013; Webb et Consi, 2001). Cette discipline vise à émuler l'intelligence à travers différents artefacts physiques. La méthodologie privilégiée pour atteindre ce but est celle de mimer les structures, fonctions et comportements des agents biologiques. Par exemple, sur le plan des structures, la bipédie est un thème exploré en robotique bio-inspirée bien que non exclusif à ce domaine, favorisant ce type de locomotion plutôt qu'un déplacement sur roue. Pour la simulation de fonctions biologiques, l'intégration et la transmission d'information à travers un RNAI illustrent aussi le domaine. Ainsi, les données entrantes provenant des capteurs de robots sont encodées numériquement et convergent vers des entités abstraites qui émulent les champs dendritiques récepteurs des RNAI. La communication entre les unités, lorsqu'elles sont en réseau, passe par des liens synaptiques, et l'émission d'un message binaire de type tout ou rien qui simule les potentiels d'action. La finalité de cette intégration trouve une correspondance en robotique par le fonctionnement d'actuateurs et la génération de mouvements ou d'actions dans le monde réel.

Dans les RNAI, les connaissances résident à l'instar de l'hypothèse naturelle dans la variabilité des poids synaptiques. L'adaptabilité provient par l'inclusion de règles variées d'apprentissage modulant les synapses. Un exemple de comportement naturel simulé en robotique bio-inspirée est celui de la coopération de robots qui, reflète le but d'accomplir une tâche commune et où souvent, celui-ci est inatteignable par un seul individu. Ainsi, le comportement social pourrait être calqué sur le monde naturel, tel qu'observé empiriquement à différents niveaux conceptuels parmi des modèles animaux.

Dans une perspective générale de robotique, les avantages à utiliser une méthodologie bio-inspirée plutôt qu'une autre dépend évidemment des objectifs. Par exemple, l'efficacité d'un système de déplacement sur roue par des robots s'avère

parfois supérieure aux agents naturels. Cependant, la locomotion sur pattes est en général plus polyvalent, entre autres lorsque les types d'environnements sont variés. Indépendamment du type de locomotion visée et sans égard à la méthodologie utilisée, l'algorithme permettant de résoudre la problématique est aussi un facteur déterminant dans l'atteinte du mimétisme. Ainsi, la robotique bio-inspirée vise une généralisation des capacités adaptatives des IA produites, telles que retrouvées dans les agents naturels. Une prémisses du domaine repose sur le fait que la nature est non seulement le meilleur exemple observé du phénomène de l'intelligence, mais de plus, il est le seul qui existe actuellement comme modèle à imiter.

#### 1.6 Agents autonomes complets virtuels et physiques

Le sujet de cette thèse concerne l'apprentissage bio-inspiré en robotique. L'application est celle des robots physiques mobiles et autonomes et vise à augmenter leurs capacités adaptatives en général. À cet effet, pour les raisons pratiques, de soucis d'économie et d'efficience, le développement des hypothèses et des tests a préalablement débuté sur des modèles virtuels. Ainsi, à l'aide de simulateurs munis d'engins de physique 3D, les résultats obtenus des expériences permettent d'approximer ceux envisagés en situations réelles et conduisent parfois même à l'émergence de nouvelles solutions.

Cependant, il n'y existe pas de meilleures représentations du monde que le monde réel lui-même, pour paraphraser Rodney Brooks. Les détails et la complexité sont forcément éludés dans l'élaboration d'un modèle d'IA et une simulation virtuelle s'avère être une approche encore plus réductionniste pouvant aiguiller les chercheurs vers des conclusions erronées.

Incidemment, l'incarnation ou la corporéité des agents cognitifs représentent un thème majeur avec lequel, les scientifiques œuvrant aux frontières des sciences cognitives et de l'IA doivent maintenant composer dans l'étude du phénomène de l'intelligence (Anderson, 2003; Brooks, 1991a; Brooks, 1991b; Clark, 1997; Mataric, 1997; Pfeifer, 2007; Riegler, 2002; Sporns, 2002; Varela *et al.* 1991; Ziemke, 2001).



Selon ces auteurs, en plus de construire des contrôleurs d'agents artificiels intelligents, leurs structures et morphologies physiques précises, tout comme le contexte environnemental dynamique, peuvent influencer la performance observée de l'acte intelligent. De leurs interactions, peuvent émerger des propriétés et des phénomènes uniques. Dans une certaine mesure, l'étude de robots physiques pourrait contribuer à une compréhension de l'intelligence par l'observation comportementale, de modèles d'IA complets, situés et incarnés. Négliger ces faits pourrait conduire vers de fausses interprétations de résultats obtenus comparativement à ceux du monde naturel, si tel était bien le but (Chandana, 2004).

Dans le développement de modèles d'IA à l'aide de simulations virtuelles, le choix des variables et conditions expérimentales s'avère crucial, car les représentations étudiées ne sont qu'une fraction d'une compréhension déjà approximative du monde réel, dans lequel a émergé l'intelligence naturelle. Dans cette perspective, l'intelligence biologique est envisagée comme un épiphénomène résultant de la triade corps, cerveau et environnement. De la réunion de ces composantes émergerait un système cognitif dynamique ne pouvant être observé que lorsque toutes les conditions y sont réunies (Chiel et Beer, 1997). Pour cette thèse, il s'agit de raisons suffisantes pour passer rapidement au développement expérimental d'une implantation physique d'une IA et la réalisation éventuelle d'une IA complète.

Indubitablement, l'agent cognitif matérialisé est situé dans un espace-temps (Steels et Brooks, 1994). Les aspects de linéarité, de direction et d'irréversibilité du temps dans un monde réel deviennent des facteurs clés de temporalité qui doivent être pris en compte dans le phénomène de l'intelligence. Par exemple, l'apprentissage, la mémoire et la causalité des actions engendrées par un agent cognitif conditionnent son futur et ne peuvent être que très pauvrement approximés par des modèles dynamiques théoriques ou simulés. L'implémentation d'algorithmes d'apprentissage dans des robots physiques devrait intégrer ces différentes notions temporelles contributives au phénomène de l'intelligence. À ce titre, la robotique mobile bio-inspirée évoluant en

milieu et temps réel peut certainement fournir des pistes de solutions dans l'étude du phénomène de l'intelligence générale.

Les agents physiques, lorsque dotés de capacités cognitives même rudimentaires, agissent et transforment leurs environnements. Par exemple, une tâche minimale d'un robot pourrait concerner des déplacements de matériel dans son environnement, jumelés à des associations perceptuelles afin de parvenir à une gestion autonome de ses besoins en énergie. Ainsi, les actions posées par le robot pourraient produire des modifications majeures de son environnement et potentiellement augmenter la complexité de l'ensemble du système cognitif observé.

Actuellement, cette seule thématique de la locomotion représente un réel défi pour un roboticien avec des résultats imprévus à anticiper avec des robots qui possèdent un nombre inférieur de degrés de liberté en comparaison avec les insectes les plus primitifs. Aussi, les roboticiens doivent composer avec une rigidité inhérente des matériaux utilisés, un temps réel à gérer ainsi que la compréhension d'un monde dynamique complexe, où il existe une non-linéarité temporelle des phénomènes courants s'y déroulant sur de multiples échelles. De plus, les déplacements dans le monde physique exigent une consommation d'énergie, ce qui implique du transport de charge supplémentaire, rendant le modèle physique lourd et imposant, en comparaison aux modèles vivants qui eux semblent être optimaux en la matière. Enfin, les divers phénomènes physiques tels que l'équilibre, l'inertie et la friction occasionnent aux chercheurs, de sérieux écueils dans leurs modèles artificiels et simulés, ce qui les oblige parfois à recourir à des solutions qui risquent de perdre leur sens au niveau comparatif de l'intelligence naturelle. Un des axes principaux de la recherche en robotique mobile se situe justement à ce niveau de maîtrise du jeu de la coordination sensori-motrice en temps réel.

En résumé, non seulement l'agent cognitif artificiel devrait être matérialisé dans un corps, se déplacer et agir dans un monde situé, mais il devra aussi être autonome. À cet égard, le terme d'autonomie est large de sens et souvent galvaudé par la communauté scientifique du domaine. De façon intuitive, l'autonomie complète d'un robot

consisterait en une absence totale d'intervention de la part des humains. Cependant, la gestion autonome du niveau d'énergie dans un environnement spécifique par des comportements exploratoires primitifs pourrait déjà être un objectif satisfaisant. Par exemple, percevoir l'endroit où recharger ses batteries en retrouvant des bornes électriques ou se diriger vers des endroits lumineux dans le cas de piles photosensibles serait des buts acceptables répondant aux besoins et à une définition minimale de l'autonomie d'un robot.

Une dernière pierre angulaire des agents cognitifs situés et incarnés est celle de la considération de l'environnement dans lequel ils évoluent. Idéalement, celui-ci devrait être le monde physique et non virtuel, celui dans lequel notre type d'intelligence a évolué et sert de base comparative. Donc, un environnement destiné aux robots devrait être non déterministe, dynamique et extrêmement versatile. L'exemple ci-haut mentionné de l'autonomie par recharge de batterie verrait le côté complexité et adaptatif rehaussés si la source d'énergie variait constamment sa localisation et sa période de disponibilité. Ainsi, la survie de ces types d'IA ainsi que leurs évolutions physiques et cognitives devraient être des préoccupations pour les chercheurs désirant réaliser des IA complètes. Bien entendu, cette thèse ne vise pas tous ces buts en robotique bio-inspirée, mais le modèle d'apprentissage est développé en tenant compte de ces diverses perspectives, évitant possiblement des pièges futurs en s'engageant dans un formalisme théorique étroit ou dans un cadre expérimental rigide.

Enfin, l'incarnation et l'aspect situé d'un ensemble d'agents cognitifs artificiels impliquent une dimension supplémentaire à retenir dans un modèle complet d'IA, soit l'impact de l'individualité et de la collectivité. Entre autres, des robots uniques et hétérogènes offriraient possiblement de meilleures capacités adaptatives, en particulier lorsque ceux-ci capitaliseraient sur des situations d'apprentissage par comportements d'imitation. D'ailleurs, une des forces de l'évolution biologique repose justement sur cette diversité, engendrée par la génétique et sur laquelle agit le mécanisme de la sélection naturelle. Une des branches de la robotique bio-inspirée s'intéresse justement à ces aspects. Cependant, l'inverse de l'individualisme semble avoir été plus attrayant

à modéliser en IA. En effet, l'intelligence collective est un des créneaux largement explorés actuellement par les chercheurs dans le domaine de la robotique cognitive et bio-inspirée.

En résumé, voici les éléments clés de cette section sur les robots physiques et les simulations virtuelles:

- La structure corporelle, la situation spatio-temporelle ainsi que l'environnement d'un agent cognitif collaborent au phénomène de l'intelligence.
- Un agent cognitif est adaptatif et innove en termes de solutions aux problèmes rencontrés.
- Un agent cognitif artificiel complet devrait être mobile et autonome, agissant et modifiant dynamiquement son environnement qui lui aussi est changeant et imprévisible.
- Le développement de modèles d'IA à l'aide de simulations virtuelles est une voie de départ raisonnable afin de tester les hypothèses visant l'implémentation de futurs modèles physiques.
- L'approche de l'IA virtuelle est limitative, réductionniste et sujette aux erreurs interprétatives en regard des conclusions obtenues sur les structures, les fonctions ou les comportements simulés. Par conséquent, l'approche de l'IA physique est la méthodologie stratégique à adopter à long terme.

### 1.7 Temporalité

Le thème de la temporalité est brièvement abordé dans cette section, car il a été le déclencheur et l'idée maîtresse derrière ce projet doctoral. L'hypothèse posée est que la temporalité représente un élément clé dans le phénomène de l'intelligence. C'est-à-

dire qu'il existerait une relation de cohérence à toutes échelles conceptuelles, de différents éléments temporels impliqués dans l'observation d'un acte intelligent produit par un agent cognitif dans un environnement physique. Donc, dans un processus de réalisation d'une IA complète, toutes les variables relatives au temps devraient être définies, correctement initialisées et implémentées, surtout en regard de leurs interrelations, particulièrement dans un contexte de temps réel.

Ainsi, le temps est une variable déterminante dans la réalisation d'un robot situé et incarné, évoluant dans un environnement dynamique. En particulier, la coïncidence des événements neuronaux est une dimension cruciale à gérer lorsque celui-ci possède un cerveau-contrôleur bio-inspiré à l'aide de RNAI. Deux hypothèses sont à valider :

***H1 : Les motifs de stimuli physiques environnementaux perçus et les réponses comportementales adaptatives secondaires des agents cognitifs représentent des événements où les relations temporelles sont causales.***

***H2 : Ces différentes relations temporelles sont particulièrement importantes dans la modélisation de fonctions d'apprentissage.***

Le temps est une propriété fondamentale de notre univers. C'est une dimension mesurable. À l'échelle non quantique, le temps s'écoule dynamiquement de façon unidirectionnelle et continue; il est particulièrement remarqué lors de changements événementiels s'opérant en temps réel. La quantification du temps dépend de l'intervalle entre les événements et de l'échelle de grandeur du phénomène observé.

Le temps peut être implicite au modèle d'étude avec un passage inhérent à la durée des événements, comme lors d'un cycle de rotation terrestre. Il peut être aussi explicite et fixe, avec une durée prédéterminée selon des paramètres telle que la milliseconde. Le temps pourrait être aussi modélisé de façon continue, comme avec des capacitances électroniques, une transmission de signaux ou bien encore discrétisé et échantillonné selon des normes fixées par un expérimentateur.

Indéniablement, les variables temporelles sont déterminantes dans l'observation du phénomène de l'intelligence. Entre autres, le processus de la mémoire implique une

gestion d'évènements chronologiques et causals se déroulant sur des échelles temporelles de durées variables. Le temps est aussi balisé pour les systèmes cognitifs naturels avec celui de leur existence. Cette fixation de la fenêtre temporelle est cruciale pour les périodes ontogéniques et phylogéniques des entités cognitives naturelles et devrait être tenue en compte dans la conception de systèmes intelligents artificiels.

Une façon d'interpréter un système cognitif dynamique (Van Gelder, 1998) est de le formaliser à travers une série d'équations différentielles décrivant la trajectoire du système selon un espace d'états et un état initial. Ainsi, il est possible de connaître les valeurs des paramètres à un temps  $t$ , d'explorer les frontières, les bifurcations, les comportements et les motifs importants du système dynamique. Les modèles théoriques dynamiques nous aident à effectuer des prédictions analytiques. Cependant, ils ne sont pas faciles d'usage lorsque le modèle comporte un grand nombre de dimensions et sont peu utiles lorsqu'il n'existe aucune solution analytique. Il convient aussi de différencier les simulations théoriques et les analyses a posteriori d'un modèle, de l'implémentation dans un robot physique d'algorithmes s'exécutant en temps réel, tel que ceux retrouvés dans cette thèse. Une alternative méthodologique doit alors être envisagée afin de résoudre ce problème d'algorithmes computationnels et de la gestion du temps réel dans la réalisation d'une IA. Spécifiquement, en ce qui concerne cette thèse, les variables concernant le temps ont toutes été discrétisées en cycles d'algorithmes.

L'hypothèse de la contiguïté temporelle est aussi une piste intéressante à investiguer (Çevik, 2014) sur ce thème. Elle sous-tend que la perception de certaines propriétés physiques de stimuli possède un impact sur le comportement des agents cognitifs. Plus précisément, c'est la distance en termes de mesure entre un stimulus et le corps qui est en relation directe avec l'organisation topologique des structures neuronales et la cohérence des mécanismes sous-jacents qui traitent le stimulus. Par exemple, un stimulus neutre répété tel qu'un flash de lumière sera traité par un mécanisme d'apprentissage non associatif de bas niveau soit, l'habituation. Cette habituation entraîne une décrémentation de la réponse comportementale au fil du temps. À



l'inverse, un stimulus tactile nociceptif tel qu'un choc électrique sera traité par le mécanisme de la sensibilisation, soit avec une réponse comportementale accentuée au fil du temps. La contiguïté temporelle dans ces phénomènes est que la vision traite des stimuli distaux au corps en premier lieu, en rapport avec le sens tactile qui lui est plus proximal. Le corrélat en apprentissage associatif tel que le conditionnement classique serait que les stimuli neutres peuvent servir une fois conditionné à prédire les stimuli inconditionnels, simplement parce que la distance temporelle et la topologie structurelle sont ainsi préalablement organisées.

En conclusion, la temporalité pourrait être considérée comme un élément clé du phénomène de l'intelligence. Il s'agit d'un sujet complexe qui mérite un vif intérêt, mais qui devra faire l'objet d'une autre étude. Avant d'entamer des hypothèses sur un lien plus direct entre la temporalité et l'intelligence, il fallait au préalable, réaliser un modèle minimal d'IA complet et bio-inspiré. L'objet de la présente thèse, soit l'élaboration de fonctions d'apprentissage pour des RNAI contrôleurs de robots, en lien avec le concept de temporalité, y procure à présent un terreau plus propice à une investigation future, dans un contexte théorique plus général.

## 1.8 Réseaux de neurones artificiels

L'ancêtre des RNA d'aujourd'hui a été créé par McCulloch et Pitts (1943), en utilisant un archétype du neurone formel. Il s'agit d'une unité mathématique abstraite possédant des propriétés particulières de calculs logiques. Ces propriétés sont exploitées par les connexionnistes afin de résoudre de multiples problèmes dans maintes sphères scientifiques, dont l'IA. Bien que plusieurs chercheurs en neurosciences et en sciences cognitives expriment leurs intuitions par des modèles connexionnistes simples, certains optent pour des modèles basés sur des mathématiques plus complexes, reflétant plus fidèlement les véritables propriétés des neurones biologiques (Hodgkin et Huxley, 1952; Koch, 1999; Van Leeuwen, 2003; Wilson, 1999). D'autres encore, s'écartent volontairement des orientations cognitives

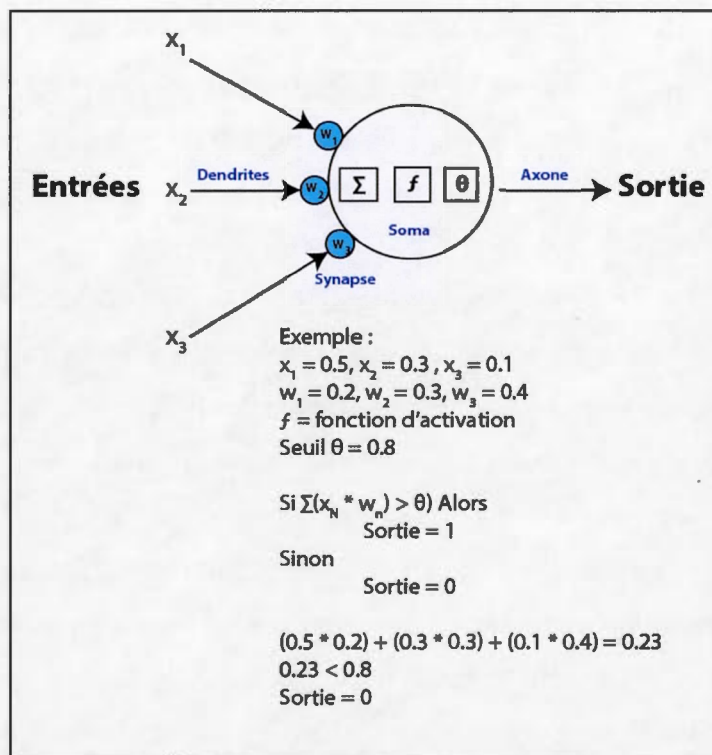


pour se concentrer sur l'essentiel des caractéristiques des fonctions mathématiques et statistiques reliées aux RNA, avec des applications et buts d'optimisation dans le champ de l'ingénierie (Dreyfus *et al.* 2002).

Dans sa forme simple (Figure 1.2), le neurone artificiel reçoit des entrées, principalement des données numériques, qu'il intègre et, selon une fonction d'activation linéaire ou non linéaire avec ou sans seuil, déclenche une réponse en sortie, habituellement de même nature. Le neurone artificiel est un modèle simple d'un neurone naturel. Minimale, le neurone artificiel accepte, intègre et transmet de l'information. Le neurone est le support, l'unité de traitement et en réseau le niveau interprétatif de l'information. Le neurone se décline en structure individuelle (représentation locale) et sous forme collective (représentation distribuée).

Mis en réseau par des liens et selon la complexité des différents paramètres reliés, il devient possible d'y observer des propriétés émergentes. Dans les réseaux à couches multiples, les neurones qui se trouvent entre les couches d'entrée et de sortie forment la ou les couches cachées. Celles-ci, couplées à des fonctions d'activation non linéaires rendent compte de plusieurs des capacités computationnelles des RNA. En contrepartie, elles en augmentent également l'opacité d'interprétation. Ainsi, les RNA se caractérisent par une architecture, une règle d'activation ou de transmission ainsi qu'une règle d'apprentissage qui module les poids synaptiques entre les neurones. Cette dernière représente un des points forts du connexionnisme et probablement le véritable pont vers les sciences cognitives (Matlin, 2005) et la modélisation de processus biologiques.

Bien que les neurones naturels soient nettement plus complexes, il existe des analogies intéressantes et certains RNA ont bonifié le neurone artificiel de base avec des paramètres additionnels. Cependant, quelle que soit la complexité des paramètres utilisés, l'amplitude du signal émis ou potentiel d'action (PA) est constante et l'intensité des stimuli entrants est intégrée par une fréquence proportionnelle des PA à la magnitude des stimuli.



**Figure 1.2** Le neurone formel représente une abstraction mathématique simpliste, mais fonctionnelle du neurone biologique. Il se caractérise par une intégration de signaux entrants, modulée par des poids synaptiques, qui suite au dépassement d'un seuil, active la production d'une réponse en sortie.

Théoriquement, les RNA fonctionnent en parallèle. Pratiquement, des unités doivent être traitées en premier dans les algorithmes, car les calculs s'effectuent par des processeurs à architecture sérielle qui n'offrent qu'une pâle imitation du parallélisme. Les RNA implémentées électroniquement en *Very Large Scale Integration* (Kier *et al.* 2006; Mehrtash *et al.* 2003) ou neuro-processeurs sont une des voies d'avenir, puisqu'ils pourraient contrer ce goulot d'étranglement inhérent aux approches classiques.

Les connaissances dans les RNA sont modélisées et distribuées à travers un réseau de poids synaptiques. Il en résulte une mémoire et une plasticité associée avec, comme incidence positive, une tolérance aux données bruitées ainsi qu'une robustesse du système. Malheureusement, la traçabilité des informations est difficile entre les entrées et sorties. Cette opacité oblige les chercheurs à une analyse a posteriori, complexe. Il

existe également des problèmes communs aux méthodes de calculs impliqués dans les RNA, particulièrement lors de la modélisation du phénomène de l'apprentissage. Entre autres, il y a le surapprentissage et la surgénéralisation, tout comme les pièges de l'atteinte des minima locaux dans l'optimisation des solutions ainsi que dans la stabilisation des RNA.

Les questions et problématiques scientifiques abordées par les RNA sont multiples. La cognition est un thème majeur, et si les problèmes visés acceptent des solutions de généralisation, d'approximation, d'optimisation ou de reconnaissances de motifs, les RNA peuvent être envisagés. En exemples, les RNA sont utiles comme outil de calibration, de prédictions financières, de reconnaissance vocale ou lors d'études comparatives d'empreintes digitales. Concrètement, les différents algorithmes ont surtout été introduits pour parfaire des systèmes d'analyses déjà existants. En sciences cognitives, les RNA sont généralement envisagés comme méthode d'approche ascendante et explicative d'un phénomène biologique spécifique à modéliser. Ils sont couramment utilisés comme outils d'apprentissage et d'optimisation dans le domaine de l'IA et de la robotique.

Le Perceptron de Rosenblatt (1957), modèle de RNA avec une architecture monocouche, a fait figure de proue dans ce domaine. Il a cependant été attaqué rapidement sur son impossibilité à résoudre le problème du « ou » exclusif, ou tout autre problème non linéairement séparable (Minsky et Papert, 1969). Les RNA multicouches et l'algorithme d'apprentissage à rétropropagation d'erreur (Bryson et Ho, 1969; Rumelhart *et al.* 1986) ont levé cet assaut majeur contre les connexionnistes en inculquant un second souffle à l'approche.

Depuis, les RNA se sont constamment améliorés et diversifiés avec l'architecture modulaire de Grossberg (1982), l'architecture récurrente avec centre attracteur de Hopfield (1982), les cartes topologiques auto-organisatrices de Kohonen (1982) et bien d'autres encore. Chacun des paramètres généraux des RNA s'est donc vu grandir en termes de variabilité, de complexité et de puissance computationnelle. Ainsi, les données d'entrée et de sortie peuvent être numériques ou non, discrètes ou continues,

uniques ou multiples, homogènes ou hétérogènes. Les architectures peuvent être monocouches ou multicouches, acycliques ou récurrentes et homogènes ou hétérogènes. Enfin, les liens les unissant peuvent être excitateurs ou inhibiteurs. La Figure 1.3 représente un exemple possible d'architecture avec les RNA.

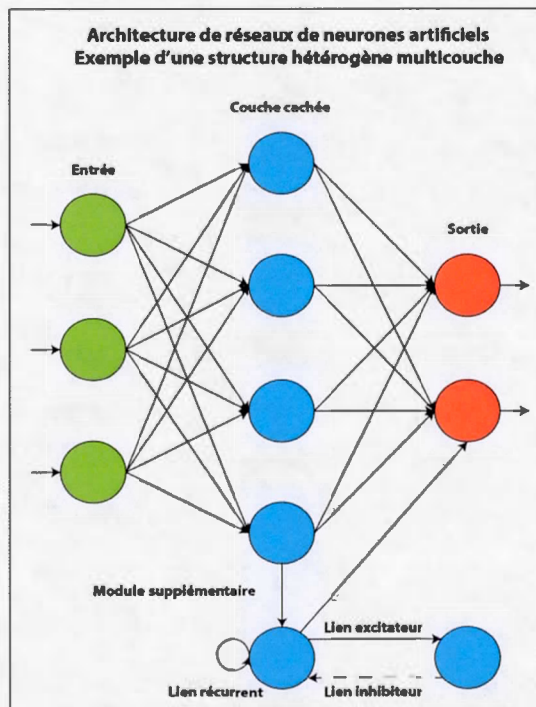


Figure 1.3 Exemple d'une architecture de RNA.

En général, un RNA utilise une règle d'activation avec seuil (p. ex. une fonction sigmoïdienne) qui, pour chaque neurone, intègre non linéairement les données entrantes pour les transformer en un signal de sortie, équivalent à un taux moyen de déclenchement des PA dans les neurones naturels. Pour les liens entre les neurones, des règles d'apprentissage modulent les poids synaptiques et reflètent l'efficacité proportionnelle des données entrantes sur la cellule de réception. Ces poids synaptiques permettent une certaine plasticité des RNA qui est à la base des propriétés des diverses formes d'apprentissage.



Il existe plusieurs types d'apprentissage, se regroupant minimalement en ceux supervisés et non supervisés. L'apprentissage supervisé implique une approximation de la solution recherchée, fournie ou non par le programmeur, où il y a correction de l'erreur en regard de l'atteinte de la solution. Pour les apprentissages non supervisés, les algorithmes doivent trouver les solutions selon des critères anticipés par l'usager. L'apprentissage peut également se faire en mode compétitif entre les neurones où selon les algorithmes, les poids synaptiques de certains neurones augmentent aux dépens de ceux des autres dans le réseau, dont l'apport est non contributif à la solution finale. Enfin, l'apprentissage de type Hebbien (associatif) propose une règle de coïncidences de PA, c'est-à-dire que si un neurone émet un PA et qu'une de ses cibles est aussi active dans une fenêtre temporelle similaire, il y aura renforcement de la synapse les unissant. Ce puissant mécanisme d'apprentissage, basé sur la plasticité synaptique activité-dépendante, est biologiquement réaliste et représente un des points d'intérêt ayant conduit à l'élaboration du sujet de la thèse.

Cependant, dans les RNA classiques, les valeurs d'entrées et de sorties sont considérées comme des moyennes reflétant la fluctuation de PA par unité de temps, en analogie à l'observation empirique de certains neurones naturels, dans des conditions méthodologiques spécifiques (périodes pouvant varier de 100 à 500 millisecondes d'intégration selon le type de phénomènes observés (Bialek et Zee, 1990). Bien que ce type de codage permette de simuler des systèmes cognitifs (Bienenstock *et al.* 1982) reposant sur des populations neuronales, cette caractéristique exclut cependant des explications relatives aux données empiriques pour des phénomènes s'observant sur une fenêtre temporelle étroite (par exemple, sous les 50 millisecondes).

De façon générale, l'obstacle majeur dans la modélisation de la cognition ou la réalisation d'une IA est de bien transposer le problème naturel étudié vers les bonnes méthodes computationnelles. Il est difficile de jauger les variables contextuelles où ces méthodes sont appliquées, tout comme l'importance des boîtes noires inhérentes aux modèles, par définition réductionnistes. En effet, il n'existe pas de mode d'emploi pour le choix d'une architecture de RNA ou la façon d'ajuster les paramètres. Le jeu d'essais

et erreurs est hasardeux pour un novice dans le domaine et rébarbatif pour un spécialiste de la méthode. Les RNA demeurent cependant de précieux outils pouvant servir autant à la compréhension de la cognition ainsi qu'à la réalisation d'une IA. Par ailleurs, des modèles hybrides existent maintenant et permettent d'automatiser certaines valeurs des paramètres. Pour les différentes raisons invoquées ci-haut, l'approche des RNA classiques et les outils algorithmiques auxiliaires ne seront pas utilisés dans cette thèse, car les objectifs fixés de cette thèse en rapport avec les contraintes temporelles nécessitent de rehausser d'un cran le niveau de bio-inspiration du modèle neuronal. Néanmoins, il aura fallu au préalable avoir une compréhension exhaustive de ce type d'outil avant de pouvoir réfuter leur utilisation dans l'élaboration des fonctions d'apprentissage pour de la robotique bio-inspirée.

### 1.9 Réseaux neuronaux artificiels dynamiques classiques

L'apprentissage de connaissances comportant des éléments temporels significatifs n'est pas nouveau. Elle requiert une forme de mémoire ayant la capacité de rappeler les événements antérieurs. NETtalk (Sejnowski et Rosenberg, 1987), un RNA de type perceptron multicouche à propagation avant et à apprentissage par rétropropagation, réussissait bien à déterminer les bons phonèmes d'une phrase, en passant un texte d'une lettre à l'autre. L'aspect temporel provenait alors du séquençage approprié des entrées fournies au RNA. Une autre façon d'envisager une persistance mnésique dans un RNA est de répliquer dans le temps l'information entre des neurones successifs, suivant une fenêtre temporelle coulissante de durée déterminée ou *Time Delay Neural Network* (Lang et Hinton, 1988). D'autres types d'architectures de RNA sont également possibles afin de réussir des apprentissages de patrons d'information codés dans le temps, comme par l'établissement de liens récurrents entre des neurones artificiels, représentant ainsi une forme de mémoire associative dynamique (Chartier et Boukadoum, 2006).



Dans les RNA à propagation avant, directe, de type perceptron multicouche, il est possible d'implémenter des mécanismes capables de transformer un RNA statique en un réseau dynamique. Ils peuvent ainsi offrir une mémoire à court terme en insérant un délai temporel synaptique, soit à la fin du réseau ou *Focused Time Lagged Feedforward Networks*, soit en distribuant le délai dans toutes les synapses ou *Distributed Time Lagged Feedforward Networks*. Il est aussi possible de trouver des équations qui permettent de résoudre une implémentation de cette mémoire dynamique, en temps continu ou en temps discret. Ces équations s'appliquent également pour des RNA à architecture bouclée ou à connexions récurrentes ou *Continuous Time Recurrent Neural Networks* (Beer, 1995). Ces derniers types de RNA sont représentatifs de modèles dynamiques ayant la capacité de simuler certains phénomènes cognitifs simples. Néanmoins, des lacunes importantes dans la constitution élémentaire de ces types de RNA limitent l'accès à des processus cognitifs plus complexes, favorisant le passage vers la dernière génération de RNA, les RNAI. Notamment, l'obtention d'une réponse de sortie à un patron d'entrée qui doit généralement passer par un grand nombre d'itérations pour la stabilisation des poids synaptiques.

#### 1.10 Réseaux neuronaux artificiels à impulsions ou pulsés

Depuis les années 1990, les RNA classiques ont vu un nouveau modèle computationnel simulant les neurones biologiques avec les RNAI (Gertsner et Kistler, 2002; Maass, 1997; Maass et Bishop; 1999; Ponulak et Kasinski, 2011). L'hypothèse théorique derrière le concept des RNAI est qu'il existe de l'information pertinente au niveau du codage unitaire des PA, dans l'ordre séquentiel des PA, dans leur coïncidence temporelle, dans leur synchronisme ainsi que dans leur intervalle de temps (Bohte, 2004). Il s'agit d'un paradigme contrastant avec les RNA classiques, qui eux fonctionnent plutôt avec des fréquences moyennes de PA représentant des données populationnelles et statiques à court terme (Rieke *et al.* 1997).

La pertinence d'introduire ce niveau de résolution temporelle dans la compréhension du traitement de l'information neuronale provient de plusieurs sources expérimentales. Par exemple, le temps de réaction de certains animaux lorsque des stimuli visuels leur sont présentés est aussi rapide que 25 msec pour une chaîne de traitements minimalement composés de quelques neurones. Aussi, la sensibilité de l'audition du hibou exige une coïncidence temporelle de l'ordre de 3 msec entre les deux aires auditives afin de bien localiser ses proies (Kempner *et al.* 1996). De plus, dans certaines conditions d'apprentissage, dont celles qui s'opèrent en un seul essai, la plupart des RNA ne conviennent pas avec leurs nombreuses d'itérations nécessaires pour stabiliser les poids synaptiques. Finalement, une des formes d'apprentissage reconnues en neurosciences est celle de la potentialisation de la mémoire à long terme (*long-term potentiation*) (LTP) (Bliss et Lømo, 1973). Un des mécanismes sous-jacents au LTP exige une coordination dans la séquence temporelle des PA pré et postsynaptique ainsi qu'une exactitude de la fenêtre temporelle de l'ordre de 50 msec (STDP). Cette coïncidence des PA possède aussi une sensibilité de moins de 10 msec des PA pour l'orientation vers une dépression ou une potentialisation de la plasticité synaptique (Markram *et al.* 1997). Ces quelques exemples marquants représentent des raisons suffisantes pour prioriser un modèle neural computationnel tel que les RNAI plutôt que la forme classique des RNA. La granularité temporelle du temps réel en robotique et dans les différentes formes d'apprentissage bio-inspirées modélisées dans cette thèse est une variable qui favorise ainsi le choix des RNAI.

Des chercheurs ont démontré de façon théorique les capacités computationnelles du codage pulsé (Benda, 2002), généré par l'information au niveau unitaire des PA. Dans une perspective appliquée et pratique, lorsqu'utilisés comme cerveau-contrôleur de robots, les RNAI semblent également mieux adaptés pour accomplir des tâches cognitives où les différentes variables reliées à la temporalité y sont prépondérantes (Brette et Guigon, 2003). Ainsi, les RNAI se distinguent de leurs prédécesseurs par un plus grand réalisme biologique tout en conservant les avantages acquis des caractéristiques des RNA traditionnels tels que la catégorisation, l'apprentissage et

l'auto-organisation (Ruf et Schmidt, 1998). En outre, il est aussi possible de simuler avec les RNAI, des phénomènes cognitifs qui n'exigent pas cette précision temporelle et ainsi calquer des résultats préalablement obtenus avec des RNA classiques.

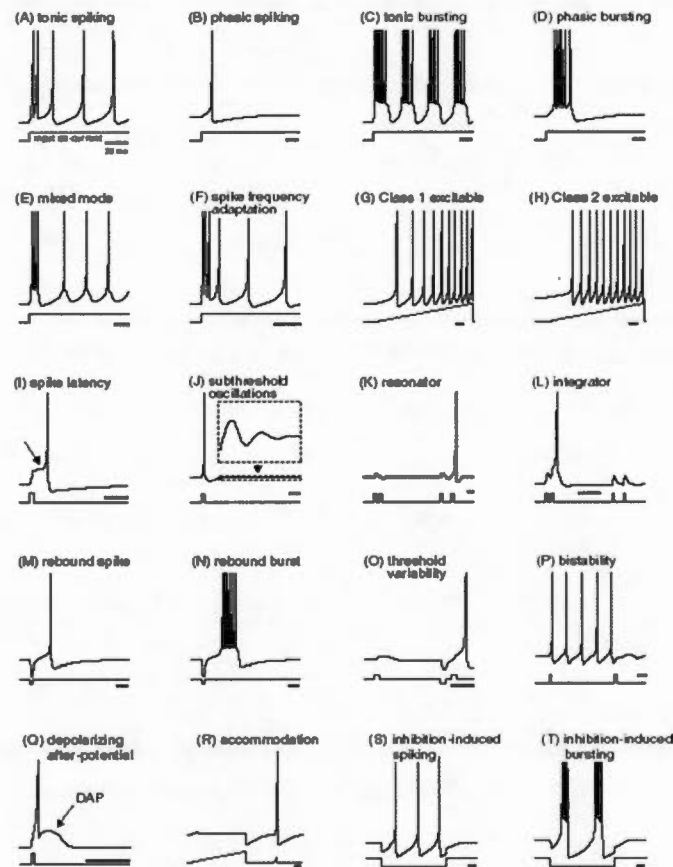
Cette différence conceptuelle de la représentation de l'information neuronale en corrélation avec le degré de précision du temps soulève un problème encore non résolu en neuroscience, soit celui de notre réelle compréhension de la transformation des signaux d'entrées et des signaux de sorties neuronales (Rieke *et al.* 1997). La problématique est qu'il existe une relation non linéaire entre les informations neuronales entrantes et la sortie dans le neurone intégrateur. Cette non-linéarité implique des variables multiples qui ne sont pas encore toutes identifiées (dépression synaptique, taux de décharge, fluctuation des ions calciques, bruit synaptique, etc.). De plus, elles possèdent des cinétiques couvrant des multiples résolutions temporelles. En conséquence, les modèles actuels de RNA et RNAI sont forcément réductifs et implémentent des équations qui approximent cette relation transformationnelle de l'information. En proportion, l'efficacité computationnelle est diminuée lorsque la précision du modèle est augmentée. Ainsi, le calcul moyen de taux de décharge par unité de temps représente le compromis fait par les RNA traditionnels, mais celui des RNAI pousse d'un cran la bio-inspiration du modèle neuronale avec le PA unitaire comme échantillonnage temporel des événements.

Ces deux derniers modèles neuronaux fonctionnent comme des neurones-points avec une absence totale de la dimension spatiale, qui pourtant offre un potentiel computationnel supplémentaire. Ainsi, les modèles de neurones artificiels à compartiments comblent cette lacune et augmentent le réalisme biologique. Cependant, à cette échelle de détails, l'efficacité d'un algorithme développé pour une entité physique artificielle complète est un objectif difficilement atteignable actuellement.

En général, la computation naturelle des événements neuronaux se traduit par l'intégration des fluctuations de courants entrants et sortants, au niveau de l'arbre dendritique (Koch, 1999). Ainsi, chacun des potentiels postsynaptiques (PPS) excitateurs (PPSE) ou inhibiteurs (PPSI) joue un rôle lors de cette intégration en faisant

varier le potentiel de membrane (PM) à tout instant. En fonction des ions et des types de canaux impliqués, la cinétique de ces PPS peut varier grandement et devenir déterminante pour l'émission de PA, particulièrement lorsque le PM est près du seuil de déclenchement (Bohte, 2007). Lorsque l'intensité et la cadence des PPS augmentent, tout comme la localisation topologique stratégique (vers le cône d'émergence où une densité de canaux ioniques prédomine), le point de non-retour est franchi, déclenchant un PA avec consommation d'énergie pour la cellule. Il s'agit donc d'un processus actif, relatif au besoin de transmission de l'information sur une longue distance, dans une entité qui est multicellulaire.

Quelquefois, il s'agit d'une salve de PA ou bouffée, déclenchée par le jeu complexe de récurrences axonales s'exprimant en de multiples configurations ou par diverses cinétiques particulières de canaux ioniques. De plus, certains neurones possèdent la propriété d'être autonomes pouvant déclencher eux-mêmes des PA par une série de cascades chimiques impliquant des canaux ioniques, protéiques et de fuites de courant, sans aucun apport de courant externe. Ces neurones deviennent ainsi des oscillateurs naturels ou intrinsèques. Il existe aussi des configurations neurales simples qui s'auto-stimulent sans réception de courant extrinsèque ou d'autonomie intrinsèque. Ces circuits peuvent conduire vers des rythmes variés de sortie de PA, comme des comportements chaotiques (Freeman, 1994) ou des générateurs de motifs centraux retrouvés, entre autres, dans plusieurs formes régulatrices de locomotions (Grillner, 1985). Avec l'ajout de la précision unitaire du PA, les RNAI conduisent à des motifs de sortie qui ressemblent de plus en plus aux multiples variations d'émissions de PA d'un neurone naturel (Figure 1.4).



**Figure 1.4** Exemples de motifs de sorties avec les RNAI. (Tiré de la fig. 1, p. 1064, Izhikevich, 2004)

Plusieurs outils logiciels sont disponibles en neuroscience. À l'aide d'algorithmes, ils permettent de modéliser le comportement des neurones naturels sur plusieurs niveaux conceptuels à la fois (de la molécule au comportement d'une entité complète). Ces simulateurs visent à mieux comprendre l'ensemble des processus neurobiologiques soutenant la cognition et le phénomène de l'intelligence (MCell, NEURON, GENESIS, Neural Network Toolbox, The MathWorks Inc.). Cependant, peu introduisent dans leurs applications une gamme étendue de faits et phénomènes neurobiologiques dans le but précis de réalisation d'une IA. Parmi ces logiciels, aucun ne pouvait établir un lien efficient entre l'édition de RNAI et leur implémentation dans des agents virtuels et physiques évoluant en temps réel. C'est pourquoi un outil adapté et spécifique à cette problématique (SIMCOG) a été développé en parallèle à cette thèse.



Dans les RNA traditionnels, les entrées et sorties sont modélisées comme des moyennes ou événements par unité de temps, car il existe une certaine relation constante entre celles-ci, ce que souligne la dualité du codage neuronal par moyenne et celui par impulsion. Cette relation a notamment été validée et observée empiriquement. Les points majeurs faisant glisser les scientifiques vers ces moyennes de PA plutôt que vers le calcul unitaire sont que l'observation des PA neuronaux ainsi que la fluctuation des PM sont variables pour un stimulus constant et ressemblent à une distribution de Poisson à ce niveau spécifique d'abstraction. Cette distribution stochastique facilite les calculs mathématiques appliqués pour des populations neuronales ainsi que pour l'analyse dynamique des différents paramètres étudiés. De plus, il existe une corrélation fonctionnelle entre des PA de moyennes et certains phénomènes neurobiologiques ou de niveaux cognitifs supérieurs.

Cependant, de plus en plus de preuves scientifiques (Abeles, 1994; Rieke *et al.* 1997) tendent à démontrer que ces événements d'impulsions neuronales unitaires sont importants à considérer à ce niveau d'échelle précis, afin de mieux comprendre et expliquer les phénomènes neurobiologiques concernant entre autres l'exactitude temporelle. Notamment, il est difficile d'expliquer et de modéliser tous les phénomènes neuronaux s'exécutant sous la fenêtre des 10 millisecondes (Gerstner *et al.* 1996). De prime abord, ces phénomènes sont plus courants que l'on imagine. Des décisions complexes telles qu'attraper une balle exige parfois des temps de réponse sous les 250 millisecondes. Dans cet exemple, la captation du stimulus débute aux couches rétiniennes, suit le trajet vers l'aire occipitale, entame un relais vers le thalamus et l'aire préfrontale, puis passe par une chaîne motrice descendante afin d'accomplir l'action appropriée. Ceci représente minimalement une bonne douzaine de synapses avec une réception, une intégration et une réponse sous les 25 millisecondes par neurone!

Un modèle intermédiaire est celui du codage neuronal par phase, particulièrement en rapport avec l'encodage dans la mémoire à court terme. Il s'agit ni plus ni moins d'une intégration de PA dans une période de temps définie dynamiquement et intrinsèquement par un ou des circuits neuronaux, lorsque l'ensemble des neurones



actifs sont en phase par des oscillations synchrones (modèle Thêta/Gamma) (Lisman et Idiart, 1995). Dans ce paradigme, le codage unitaire précis devient moins important à cette fenêtre de lecture. L'aspect critique du codage et du décodage réfère plutôt aux cycles phasiques entre neurones ou populations neuronales et qui soulignent la dualité onde et impulsion neuronale. Cette cohérence émergente est cependant envisagée et étudiée empiriquement sur des populations neuronales massives. Il s'agit d'une perspective intéressante, mais se situant à une autre échelle de modélisation.

Globalement, les phénomènes neuronaux significatifs à réponse rapide peuvent être classés en quatre catégories. Il s'agit de la détermination d'un évènement sur un premier PA, la coïncidence de deux PA comme évènement décisionnel, la synchronisation rapide entre deux neurones ainsi que l'initialisation des PA (Naundorf *et al.* 2006) comme avec les oscillateurs intrinsèques. Il est aussi possible d'inclure les processus d'apprentissage selon la théorie de Hebb (1949). L'hypothèse de cette théorie est qu'il y a renforcement de la synapse entre le neurone présynaptique et le neurone postsynaptique, où l'exactitude de la fenêtre temporelle par laquelle ces deux évènements se succèdent, ce qui contribue à modifier les synapses impliquées. Selon certaines études, un mince décalage séquentiel de 20 millisecondes renforcerait au maximum une synapse, avec un effet de courbe en cloche pour des temps s'éloignant de ces valeurs. À l'inverse, il y aurait une diminution de l'efficacité de la synapse pour les PA asynchrones ou encore, lors des situations où l'élément postsynaptique précède la décharge du neurone présynaptique (Song *et al.* 2000; Roberts et Bell, 2002).

De plus, l'exactitude temporelle ou le délai à respecter lors d'une association de type récompense avec ou sans stimulus inconditionnel sont des caractéristiques majeures des modèles computationnels de ce phénomène ou *Temporal Difference Learning* (Alexander, 2007; Suri, 2002; Suri et Schultz, 1999; Sutton et Barto, 1990). En analogie, les neurones naturels de type dopaminergique semblent potentialiser une association favorable entre deux évènements séquentiels significatifs pour l'entité et deviennent ainsi un rouage supplémentaire dans le processus de l'apprentissage. Cette

modulation favorable au renforcement synaptique pourrait impliquer un processus de gestion à fine granularité temporelle pour bien en assurer la fonctionnalité.

Il existe plusieurs modèles décrivant les aspects dynamiques principaux des synapses (Natschläger *et al.* 2001; Poznanski, 1999). À cause du niveau de détails et du coût computationnel, peu de chercheurs utilisent ceux-ci conjointement avec des implémentations visant des concepts cognitifs de haut niveau ou vers une plate-forme physique, tel qu'un robot. L'efficacité des algorithmes utilisés et les caractéristiques visées doivent donc être choisies adéquatement. Par exemple, les PA une fois déclenchés ne produisent une réponse postsynaptique que dans 10 à 30% des cas. Également, l'effet de renforcement de la synapse qui, dans un temps court, peut se modifier de plusieurs centaines de pourcents en efficacité, et ce à des échelles temporelles variables. Enfin, il y a aussi les différents phénomènes de base de la plasticité, du remodelage et de l'homéostasie synaptiques qu'il faudrait tenir en compte dans un modèle synaptique plus complet.

Les situations mentionnées ci-haut, exigent un codage et décodage de l'information au niveau temporel unitaire des PA, a contrario des RNA classiques qui fonctionnent par codage de moyennes de PA. (Gertsner et Kistler, 2002; Izhikevich, 2003; Maass et Bishop, 1999; Rieke *et al.* 1997; Wilson, 1999). Dépendamment des questions posées au préalable et du niveau de modélisation recherchée, les codages unitaires, de moyenne ou de phase ne sont pas totalement mutuellement exclusifs. Autant au niveau théorique que sur le plan fonctionnel, une modélisation d'entité cognitive artificielle devrait intégrer les informations disponibles sur des échelles temporelles variables.

Une caractéristique essentielle des RNAI est celle d'être représentée par un état interne, mimant le potentiel de membrane (PM) des neurones biologiques. Cette variable à base de conductance électrique peut être affectée dynamiquement par les différentes entrées simulées de courants positifs et négatifs. Le neurone artificiel possède ainsi une différence de potentiel entre le milieu intérieur et extérieur de la cellule. Ainsi, une variation de voltage dans le milieu intracellulaire, lorsque perturbé par des entrées et sorties de courants, entraîne le PM vers une autre variable essentielle

des RNAI, le seuil. Lorsque celui est atteint, l'état suivant du neurone atteint inévitablement celui d'une émission de PA.

Incidemment, il existe plusieurs types de RNAI. Le *leaky integrate-and-fire* (LIF) et le modèle général de LIF ou *spike response model* (SRM) en sont deux exemples classiques. Des modèles biophysiques plus complexes existent également tel que les abstractions neuronales à compartiments et ceux des modèles à canaux ioniques tel que le modèle de Hodgkin et Huxley (1952). Il devient évident que plus le niveau de détails simulé est important, plus il en sera coûteux au niveau computationnel.

Le modèle de RNAI est donc intéressant pour ce doctorat, car il permet intrinsèquement de coder l'information avec la résolution temporelle d'un seul PA comme évènement phénoménologique significatif. Cependant, les modèles de RNAI utilisés actuellement requièrent aussi de grandes capacités computationnelles (Brette, 2004). Toutefois, certains modèles sont potentiellement exécutables dans des robots physiques entièrement autonomes (Christodoulou *et al.* 2002).

Le choix du modèle neuronal bio-inspiré utilisé a donc été crucial pour cette thèse ayant comme sujet l'apprentissage en neuro-robotique. Les caractéristiques retenues étant d'être capable de supporter le temps réel pour l'implémentation dans un robot physique ainsi que celle du niveau de détail temporel de l'ordre de la durée d'un PA, nécessaire pour les associations neuronales et les changements synaptiques consécutifs. Le nombre de paramètres du RNAI choisi ainsi que la taille maximale du RNAI ont également été des éléments essentiels à considérer dans le choix du modèle neuronal afin de ne pas excéder les limites et capacités computationnelles des ordinateurs lors des simulations et des robots physiques.

En somme, cette thèse défend la possibilité de réaliser plusieurs fonctions d'apprentissage primaires dans des RNAI, résolues par des mécanismes de plasticité synaptique pour un contexte d'application robotique. L'hypothèse est soutenue par des résultats comportementaux similaires au monde naturel, pour des robots simples possédant quelques senseurs et effecteurs et comptant tout au plus, une cinquantaine de neurones et une centaine de synapses. Cette perspective se veut un calque du cerveau

des animaux invertébrés les plus primitifs. La stratégie employée a été d'adapter les caractéristiques de base des RNAI dans un algorithme efficient, capable d'effectuer un cycle dans un intervalle de temps près de la milliseconde et dans un environnement physique réel (Brette et Guigon, 2003; Izhikevich, 2003). Au modèle de RNAI choisi a été incorporée de façon cohérente une des trois fonctions d'apprentissage primaires bio-inspirées et incrémentales en complexité, soit l'habituation, le CC et le CO. La modélisation de ces règles d'apprentissage reflète un large éventail de phénomènes neurobiologiques avec production de résultats adaptatifs du niveau de la synapse au niveau comportemental d'entités virtuelles et physiques.

En résumé sur les RNA classiques, dynamiques et les RNAI:

- Le connexionnisme est un courant de pensée qui tente d'expliquer la cognition et l'intelligence par un modèle conceptuel analogue aux neurones naturels.
- Le neurone artificiel est un modèle abstrait pouvant être connecté en réseau, possédant des propriétés émergentes relatives à la complexité de l'architecture.
- Les RNA fonctionnent théoriquement en parallèle avec une réception simple ou multiple, une intégration, un traitement et une émission du type tout-ou-rien de l'information.
- La connaissance des RNA est distribuée à travers des poids synaptiques.
- Les RNA sont capables de simuler plusieurs types d'apprentissage dynamique, supervisé et non supervisé.
- Les RNA dynamiques s'avèrent être un bon choix comme contrôleurs d'agents cognitifs pour l'adaptation aux situations se déroulant en temps réel.



- Les RNAI représentent une alternative biologiquement plus plausible que les RNA avec un modèle de codage de l'information au niveau unitaire du potentiel d'action.
- Les RNAI possèdent des caractéristiques temporelles uniques qui les favorisent dans le développement de fonctions d'apprentissage pour la robotique bio-inspirée, notamment les coïncidences de PA et les temps d'intervalles.

### 1.11 Cognition et apprentissage en robotique

Poursuivre le but de réaliser des IA dont les capacités cognitives simuleraient les capacités humaines représente un défi de taille pour les chercheurs d'aujourd'hui. Malgré l'argumentation philosophique de certains qui anticipe l'échec inévitable du projet, plusieurs scientifiques persistent à vouloir résoudre les mystères de l'intelligence, pendant que d'autres s'engagent sur la voie de reproduire les caractéristiques principales dans des artefacts virtuels et réels. Cette thèse doctorale s'inscrit dans cette dernière démarche et apporte une contribution au domaine.

Une des prémisses en IA est qu'il existe des composantes fonctionnelles élémentaires définissant l'intelligence biologique, identifiables et transférables vers des substrats non biologiques. À cet égard, Il convient de scinder les différents aspects du phénomène de l'intelligence naturelle afin de cibler des éléments précis pour ensuite échafauder un modèle artificiel. Bien que les définitions de l'intelligence et de la cognition divergent dans leurs concepts sous-jacents, la communauté scientifique s'entend toutefois sur le point que la possession des capacités adaptatives d'apprentissage représente un réel attribut du phénomène de l'intelligence et, par extension, de la cognition. Tel est le sujet spécifique de ce doctorat, visant à explorer le thème de l'apprentissage en robotique à travers une perspective de bio-inspiration.

Cette exploration dépasse le cadre de l'étude théorique du phénomène, puisque les règles d'apprentissage sont mises de l'avant et vérifiées dans des robots physiques. La

démarche est de reproduire un certain nombre de phénomènes de l'apprentissage naturel, puis d'équiper les robots de ces modèles. Ainsi, dans ce paradigme de neuro-robotique et du cadre de trois règles d'apprentissage primaire, cette thèse démontre que les robots apprennent par des mécanismes internes similaires aux agents biologiques, soit par une modulation de la plasticité synaptique.

À ce stade, il convient de définir plus précisément l'apprentissage. De façon simple et acceptable, l'apprentissage représente un processus de changement de la connaissance qui résulte de l'expérience (Mazur, 2013). Il s'agit cependant d'une capacité adaptative naturelle complexe, composée de plusieurs mécanismes et structures. La mémoire représente un de ces concepts clés sous-jacents au phénomène de l'apprentissage et qui est encore non résolu. La mémoire implique minimalement un support physique des informations et un mécanisme d'encodage lors de la phase d'acquisition. Un processus différent est tout aussi essentiel pour retrouver cette information et enfin, un mécanisme supplémentaire est nécessaire pour modifier l'information dans les cas d'oublis et des apprentissages. Accessoirement, le processus de la mémoire semble devoir recourir également à des activités de maintenance pour parer à la dégradation de l'information dans le temps.

Dans le domaine de l'apprentissage, les neurosciences et la psychologie fournissent une quantité impressionnante de données empiriques, aidant à la compréhension globale du phénomène. C'est en regard de certaines de ces récentes avancées ou inspirations biologiques que les mécanismes d'apprentissage pour les robots sont élaborés dans cette thèse. L'actuelle compréhension de la dynamique des cellules neuronales impliquées dans l'apprentissage est telle qu'il est maintenant permis d'échafauder des modèles d'analyses computationnelles complexes (Baxter *et al.* 2013), tel que des contrôleurs de robot de type RNAI et des fonctions synaptiques dynamiques, mimant la capacité adaptative des phénomènes naturels.

Incidemment, l'apprentissage est un thème large, incluant plusieurs sous-types déclinés à divers niveaux conceptuels. Au bas de l'échelle se trouve l'adaptation sensorielle. Il s'agit d'une diminution de la réponse à des stimuli constants par une

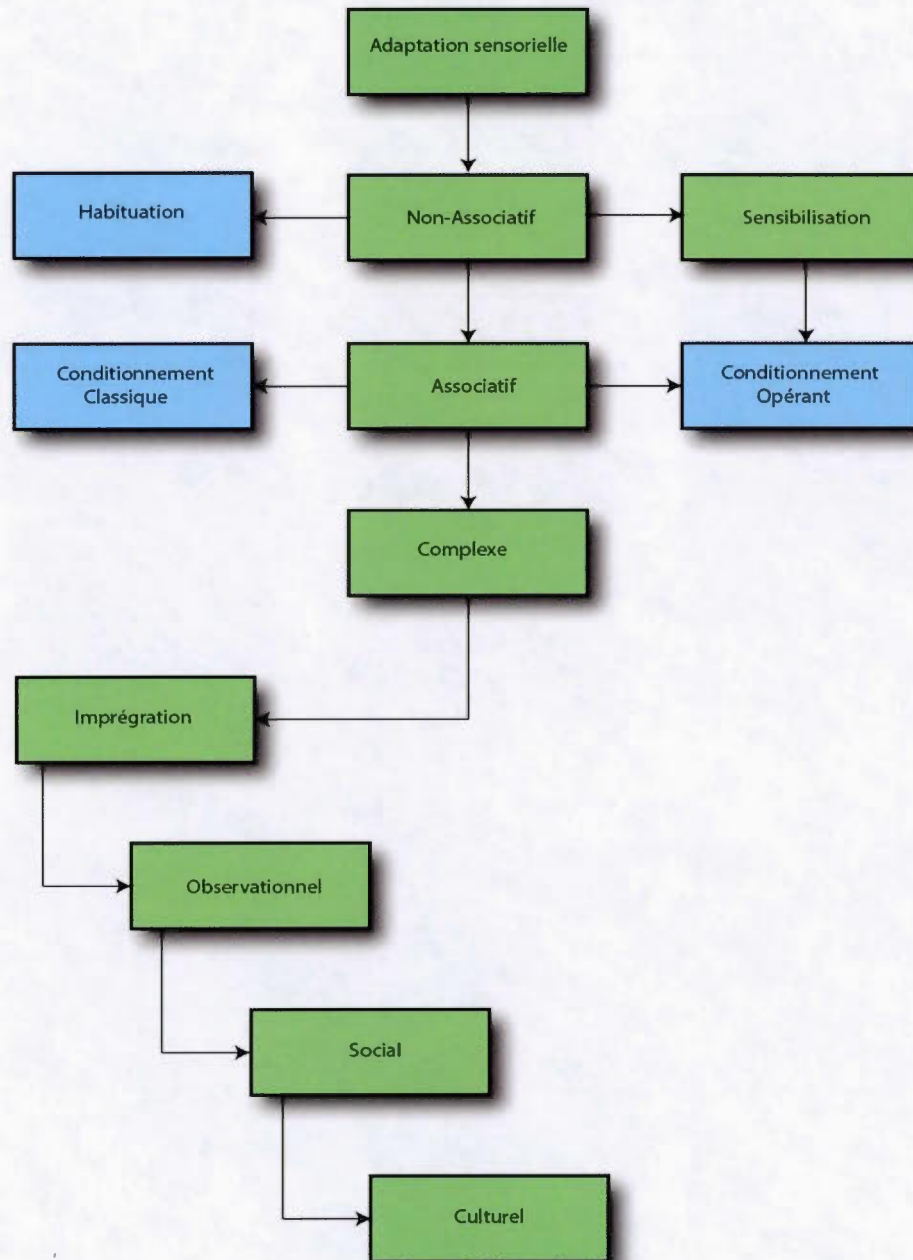


usure locale des structures et en rapport avec les cinétiques des cascades biochimiques. Une autre catégorie est celle des apprentissages non associatifs avec l'habituation et la sensibilisation. Ces phénomènes diffèrent de l'adaptation sensorielle par différentes formes de mémoires prolongées des expériences passées et répétées. Ces changements se situent majoritairement au niveau présynaptique. L'habituation est un phénomène modélisé pour cette thèse.

La catégorie suivante d'apprentissage est celle des associations, soit entre deux stimuli pour le conditionnement classique ou par le conditionnement opérant entre une action et un stimulus. Ces deux derniers phénomènes sont également modélisés dans cette thèse et présentent une complexité incrémentale en rapport avec l'habituation, soit d'inclure un élément neuronal postsynaptique et une règle d'apprentissage modifiant les poids synaptiques.

Des types d'apprentissage de plus haut niveau, tels que l'imprégnation, l'imitation par observation, l'apprentissage social et culturel, requièrent probablement des structures et mécanismes plus complexes qui ne seront pas abordés dans ce doctorat. La Figure 1.5 résume les types d'apprentissages.

## APPRENTISSAGE



**Figure 1.5** Diagramme représentant de façon incrémentale, la complexité des divers types d'apprentissage.



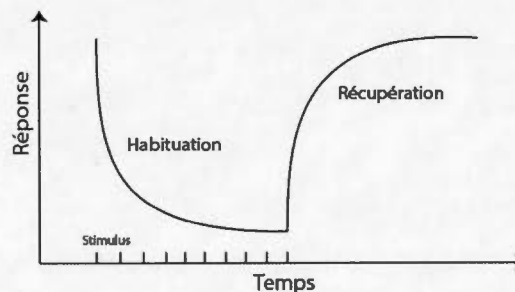
## CHAPITRE II

### HABITUATION

#### 2.1 Volet neuroscience

L'habituation est une forme d'apprentissage élémentaire qui est conservée phylogénétiquement depuis les modèles animaux à système nerveux simple comme le ver *Caenorhabditis elegans* jusqu'à l'homme. Le caractère universel de l'habituation souligne son aspect critique dans la survie des espèces. Bien qu'il s'agisse d'un type primaire d'apprentissage, les mécanismes de l'habituation sont complexes et soulèvent encore bien des questions de la part des chercheurs.

L'habituation est une forme d'apprentissage dite non associative, car la réponse comportementale ne dépend que d'un élément, les stimuli entrants. La propriété majeure de cette réponse, lorsque des stimuli sont répétés, est que la fréquence des comportements consécutifs s'adapte selon une cinétique exponentielle décroissante (Castelluci *et al.* 1974). Un deuxième élément caractérisant l'habituation est celui de la phase de récupération spontanée lorsque les stimuli cessent, graphiquement représenté par une courbe exponentielle croissante et qui n'est pas simplement l'inverse de l'habituation (Figure 2.1).



**Figure 2.1** Le phénomène de l'habituation décrit une diminution comportementale de la réponse à un stimulus répété. Lorsque les stimuli cessent, une phase de récupération débute, décrivant une fonction qui n'est pas strictement l'opposé de celle de la phase précédente de décrémentation.

L'intérêt d'étudier cette fonction est multiple, mais repose avant tout sur une meilleure compréhension du processus général de la mémoire et des bases biologiques communes aux diverses formes d'apprentissage. L'étude de l'habituation s'avère avantageuse, car selon les différentes techniques utilisées, ce processus d'apprentissage est observable du niveau moléculaire et génétique, jusqu'au niveau de son expression comportementale (Rose et Rankin, 2001; Dong et Clayton, 2009; Giles et Rankin, 2009). Outre le résultat évolutif, les fonctions apparentes de l'habituation semblent être l'économie d'énergie face aux stimuli répétés, le filtrage d'information non pertinente, la suppression de comportements non souhaitable, mais aussi, le rehaussement de stimuli nouveaux. Cette dernière fonction est possiblement une étape précurseur au phénomène de l'attention (Balkanius, 2000).

L'habituation est définie comme un processus d'apprentissage. Suite à l'expérience de la répétition de stimuli similaires, divers changements neuronaux sont induits et produisent une atténuation durable des réponses comportementales. Cette réponse est observable sur plusieurs niveaux d'abstraction, variant d'une légère diminution d'amplitude d'un PPS pour une synapse à celui de l'arrêt complet d'une action complexe impliquant la dynamique de milliers de neurones. Précisément, l'habituation est un processus qui est soutenu par des phénomènes de plasticité synaptique en majorité par une dépression homosynaptique (Stopfer et Carew, 1996) induite par des stimuli sensoriels répétés.

L'étude de l'habituation a débuté en 1956 (Thorpe, 1956), mais ce n'est qu'à la fin des années 60 (Thompson et Spencer, 1966; Groves et Thompson, 1970) que s'établit un modèle théorique défini par une série de caractéristiques. Récemment, des chercheurs du domaine (Rankin *et al.* 2009) ont actualisé ces critères. Le modèle d'habituation développé dans le cadre de cette thèse doctorale est fortement inspiré de cette dernière liste de propriétés (voir plus bas le résumé des caractéristiques).

Dans un premier temps, l'apprentissage par habituation a d'abord été distingué du mécanisme d'adaptation sensorielle qui, a priori, aboutit à la même finalité, soit de limiter l'effet de stimuli répétés. Par exemple, en réponse à une petite frappe (1 newton)

sur un plat de pétri contenant des vers nématodes *C. elegans* nageant sur une surface de Agar, ceux-ci arrêtent brusquement leur mouvement avant, en effectuant un bref mouvement arrière, approximativement de longueur équivalente à celle de l'animal, soit près de 1 millimètre (Rankin *et al.* 1990).

Ce phénomène est une réponse comportementale mécanosensorielle non localisée et réversible qui est équivalente à celle de la surprise chez les animaux supérieurs. Le circuit neuronal de *C. elegans* impliqué dans cette habitude à la stimulation mécanosensorielle répétée (Kitamura *et al.* 2001), est complètement élucidé (Chalfie *et al.* 1985; Wicks et Rankin, 1995). Pour une multitude de raisons, cet invertébré est un modèle de choix pour l'étude de l'habitude ainsi que pour tous les autres types d'apprentissage (pour un article de revue, voir Ardiel et Rankin, 2010).

En produisant des frappes répétées dans une même session d'apprentissage, un certain nombre de fois avec un intervalle interstimulus défini (IIS: secondes à minutes) et en mesurant les mouvements relatifs de recul des vers, des chercheurs ont observé que ce comportement décroît progressivement avec le temps (Rankin *et al.* 1990). Ce paradigme expérimental standard pour *C. elegans* a permis d'observer la caractéristique essentielle de l'habitude, soit une diminution de la réponse comportementale face à des stimuli répétés. Cette décrémentation comportementale en amplitude et en fréquence de la réponse demeure stable pour un certain temps, qualifiée comme une mémoire à court terme (minutes à heures) des stimuli antérieurs. Cependant, une fois l'habitude maximale complétée, l'application d'un faible choc électrique (autre type de stimulus) sur le pétri, résulte de nouveau en une observation du comportement de recul complet, éliminant la causalité du processus d'adaptation sensorielle, qui repose sur une fatigue du système support et qui nécessite une période de repos pour le recouvrement.

Ce phénomène nommé déshabitude, observé par l'ajout d'un stimulus nouveau, différent ou fort, est encore mal compris au niveau moléculaire (Steiner et Barry, 2014). C'est une caractéristique importante de l'habitude que d'être spécifique à un type de stimulus. Cette spécificité peut être parfois généralisable à d'autres stimuli lorsqu'ils



sont dans une même modalité sensorielle, lié probablement à des préférences neuronales génétiques (Pilz *et al.* 2014). Par exemple chez la souris, l'habituation est maximale pour des stimuli sonores répétés avec une tonalité pure et une fréquence fixe de 14 kHz. Ce qui correspond à la captation optimale dans cette modalité sensorielle chez cet animal.

Les chercheurs ont aussi trouvé une relation directe entre la fréquence des IIS et la réponse comportementale de l'habituation. Plus le IIS est court, plus l'habituation est rapide et inversement (Davis, 1970; Broster et Rankin, 1994). Cette autre caractéristique a contribué à révéler qu'il existe une hiérarchie de mémoire reliée à l'habituation (Steidl *et al.* 2003), classifiée selon leur durée d'action allant de quelques minutes à plusieurs jours (Beck et Rankin, 1995). Incidemment, un IIS court induit une mémoire à court terme et un IIS long conduit à une mémoire à long terme de l'habituation.

Plusieurs autres types de mémoires intermédiaires (Sutton *et al.* 2001) ont été observés, soutenus par des processus biophysiques et génétiques distincts. Majoritairement, les divers processus liés à l'habituation reposent sur une dépression synaptique activité-dépendante (Byrne, 1982) ainsi que sur une modulation de l'excitabilité neuronale. La morphologie de l'élément présynaptique (nombre/taille des varicosités et des vésicules contenant des neurotransmetteurs) (Rose *et al.* 2002), le nombre de canaux ioniques  $K^+$  et  $Ca^{++}$  (Cai *et al.* 2009) ainsi que le nombre et le type de récepteurs postsynaptiques (ratio NMDA/AMPA et types de sous-unités pour le processus de mémoire à long terme, Rose *et al.* 2003; Glanzman, 2010), représentent autant de variables mesurées à la baisse pendant l'habituation.

Autre variable cruciale, la vitesse de récupération post-habituation est inversement proportionnelle à la fréquence des IIS. Cette fréquence s'avère donc être un bon prédicteur de la vitesse de récupération, bien plus précis que la mesure du taux courant d'habituation, ou encore le nombre absolu de stimuli répétés (Rankin et Broster, 1992). Cette autre caractéristique essentielle de l'habituation a ainsi conduit les chercheurs à former l'hypothèse que les différents types de mémoires de l'habituation reposent entre

autres sur des mécanismes biophysiques qui sont IIS-dépendants. Ainsi, selon les différents patrons de IIS, les changements synaptiques consécutifs proviennent de processus biochimiques spécifiques, en lien ou non avec la transcription d'ARN et la synthèse protéique (Beck et Rankin, 1995; Bailey *et al.* 1996; Nuttley *et al.* 2001; Ezzeddine et Glanzman, 2003; McNamara *et al.* 2008).

En d'autres termes, la cellule neuronale semble encoder la fréquence des stimuli entrants en empruntant des voies diverses de changements synaptiques, reflétant des catégories déterminées de IIS qui conduisent à autant de types différents observables du phénomène de cette mémoire non associative. L'élaboration du modèle d'habituation de cette thèse a précisément capitalisé sur cette hypothèse et des propriétés temporelles que sont les durées variables des périodes d'habituation et de récupération en lien avec des catégories distinctes d'IIS. Aucun algorithme modélisant l'habituation pour des RNAI n'avait encore intégré autant de ces paramètres temporels, corroborant les études expérimentales actuelles du domaine.

Un autre modèle d'invertébré ayant contribué aux découvertes reliées à l'habituation est celui du mollusque marin *Aplysia californica* (Pinsker *et al.* 1970; Byrne, 1982; Hawkins *et al.* 2006). Par exemple, l'habituation du réflexe défensif de retrait du siphon de l'Aplysie, lorsque stimulé répétitivement par une force mécanique, est démontrée aussi du niveau cellulaire au niveau comportemental. Les chercheurs tentent de résoudre le processus complet chez cet invertébré possédant un cerveau de plus de 20 000 neurones, en investiguant aussi l'habituation au niveau moléculaire et génétique.

Il a été démontré que le réflexe de retrait du siphon est observable et modulable par de l'habituation en stimulant n'importe quels circuits afférents sensori-moteurs, tel que celui la queue de l'Aplysie. Cette observation permet de déplacer le lieu focal de l'habituation, de la synapse usuelle sensori-motrice impliquée pour le siphon, vers des interneurones de deuxième ordre, puisque dans ce dernier réflexe (stimulation mécanique de la queue et retrait du siphon), il n'existe pas de lien synaptique direct entre la queue et le siphon. À cet égard, des chercheurs (Bristol et Carew, 2005) ont mis en évidence le rôle clé de neurones inhibiteurs (Krasne et Teshiba, 1995) dans la

modulation mono et postsynaptique par une architecture de neurones de premier et deuxième ordre. Dans ces circuits, le rôle des neurones inhibiteurs est interprété comme limitant l'amplitude des PPS, soit directement aux motoneurones impliqués ou indirectement par d'autres microcircuits, avec ou sans ajout d'habituation. Cet axe de recherche est important dans l'élaboration d'un modèle computationnel complet de l'habituation naturelle, dépassant simplement le niveau de la fonction à une seule synapse sensori-motrice pour migrer vers une compréhension globale de la diminution comportementale au niveau d'un circuit.

Les conditions expérimentales retrouvées dans les protocoles varient beaucoup dans la littérature sur l'habituation, selon le type de stimulus, l'animal, l'espèce (c.-à-d. souris vs rat), la modalité sensorielle et le but de l'observation. En particulier, il existe une bonne variance dans l'application des patrons temporels de stimuli (IIS), allant de la milliseconde aux journées! Par exemple, une frappe mécanique de 30 stimuli répétés à 10 secondes d'intervalle est un protocole d'habituation standard pour *C. elegans* afin d'observer une mémoire à court terme, pour une récupération complète en 10 minutes du comportement atténué de recul. Si la fréquence des IIS est augmentée à 60 secondes, la récupération complète se fait alors en 50 minutes. À noter qu'afin de tester le taux d'apprentissage par habituation d'une série de stimuli, une mesure indirecte du taux de récupération est effectuée à des temps variés après les derniers stimuli, tel que 30 secondes, 10 minutes et 20 minutes pour un protocole standard avec *C. elegans*. Une mesure de la réponse comportementale pendant l'entraînement révèle le niveau d'habituation et les mesures tests après l'entraînement, celui de la récupération.

Pour l'Aplysie, un protocole de 10 frappes mécaniques d'une durée de 25 millisecondes, appliqué à 10 secondes d'intervalle sur la queue de l'Aplysie entraîne un processus d'habituation du réflexe de retrait du siphon sur une période d'environ 5 minutes (mesure indirecte des PPS induits sur le motoneurone receveur du siphon) (Bristol et Carew, 2005).

Dans le système olfactif chez la souris, 5 applications d'une odeur pendant 20 secondes avec un IIS de 10 secondes résultent en une récupération complète en moins

de 10 minutes pour la mémoire à court terme. Lorsque la durée est augmentée à 50 secondes et l'IIS à 5 minutes, la récupération complète prend un peu plus de 30 minutes pour la mémoire à long terme (McNamara *et al.* 2008). À noter que dans cette dernière étude, l'engagement vers la voie de la mémoire à long terme selon le type de protocole utilisé, détermine aussi le degré de généralisation possible des types d'odeurs avec comme résultat, qu'il n'y en a aucune à court terme et un maximum avec la mémoire à long terme.

Dans d'autres protocoles, la répétition de blocs de stimulation avec des intervalles de temps définis produit aussi une mémoire à long terme de l'habituation. Pour obtenir une mémorisation à long terme, les essais pendant les entraînements doivent être distribués ou espacés dans le temps selon des intervalles interbloc (IIB) et non appliqués en masse (Beck et Rankin, 1997; Masini *et al.* 2008).

Ces fenêtres de temps d'apprentissage à IIB sont aussi variables selon les protocoles utilisés. Avec l'exemple de *C. elegans*, trois blocs séparés par une heure de repos, où 20 stimuli sont répétés à 60 secondes d'intervalle, conduisent à une rétention à long terme de l'information sur plus de 24 heures (Rose *et al.* 2002). Ces blocs de stimulations provoquent une accumulation graduelle de la mémoire ou potentialisation de l'habituation. Cette voie de mémoire à long terme de l'habituation semble être reliée à une plus grande efficacité synaptique du système glutamate et des récepteurs NMDA/AMPA. Cette efficacité dépend aussi du temps entre les blocs ainsi que du nombre de blocs d'entraînement.

Différentes autres propriétés définissent aussi l'habituation. L'effet de l'amplitude et la durée d'application des stimuli, le type de stimulus, l'effet ontologique (Rankin et Carew, 1987), la modulation par des neurotransmetteurs (Hills *et al.* 2004; Kindt *et al.* 2007), la comparaison sur différents sites d'un même système sensoriel (Frost *et al.* 1997), la comparaison sur différents systèmes sensoriels et différents systèmes nerveux animaux, sont autant d'éléments étudiés qui caractérisent le phénomène.

Enfin, même si l'habituation est une forme d'apprentissage spécifique aux stimuli (Vogel et Wagner, 2005), le phénomène semble généralisable en intramodal avec des



exceptions (Pilz *et al.* 2014). L'habituation est aussi démontrée comme étant dépendante du contexte (Rankin, 2000). En effet, les expérimentations sur *C. elegans* faites en présence de bactéries *E. coli*, montre que l'habituation devient plus rapide si l'expérience est répétée sans les bactéries. Donc, l'habituation semble aussi interagir avec les autres mécanismes d'apprentissages qui sont de types associatifs (Kindt *et al.* 2007). En résumé, les critères principaux de l'habituation sont:

- 1- Réponse diminuée en fréquence et/ou magnitude et/ou en durée lorsque des stimuli non nociceptifs sont répétés.
- 2- Récupération spontanée complète (échelle temporelle variable) de la réponse habituée lorsque les stimuli cessent.
- 3- Potentialisation de l'habituation secondaire aux expériences d'apprentissage selon les IIS/IIB.
- 4- Modulation de la cinétique de l'habituation selon la variabilité des IIS.
  - a. Plus le IIS est court, plus l'habituation et la récupération sont rapides, fortes, complètes et inversement.
- 5- Modulation de la cinétique de l'habituation selon une variabilité d'amplitude des stimuli dans une même modalité.
  - a. Plus le stimulus est faible, plus l'habituation est forte et inversement.
- 6- Effet sur le début du temps de la récupération spontanée lorsque les stimuli persistent au-delà de l'asymptote.
- 7- Spécificité intrasensorielle et de site dans l'habituation aux stimuli avec possibilité de généralisation.
- 8- Augmentation d'une réponse habituée lorsqu'un stimulus différent ou fort est appliqué (déshabituation).
- 9- Une habituation de la déshabituation est possible lorsqu'un stimulus différent ou fort est répété.
- 10- Certaines variables temporelles (IIS/IIB) utilisées dans les divers protocoles expérimentaux conduisent à une mémoire à long terme de l'habituation.
  - a. Plus les intervalles sont longs, plus la rétention de l'habituation est forte et longue.

## 2.2 Volet IA

Dans la littérature en IA, l'adaptation de modèles théoriques du phénomène de l'habituation biologique n'a pas tardé à suivre avec l'apparition des premiers algorithmes computationnels capables de mimer la dynamique de l'habituation. À cet

égard, Stanley (1976) a produit un premier modèle d'habituation à l'aide d'une simple équation différentielle de premier ordre. D'autres chercheurs ont emboîté le pas (Innis et Staddon, 1989; Sutton et Barto, 1990; Wang, 1994; Staddon et Higa, 1996; Staddon *et al.* 2002; Alonso *et al.* 2005; Staddon, 2005), avec des modèles computationnels de l'habituation mathématiquement de plus en plus complexe incluant un nombre grandissant des caractéristiques originellement proposées par Thompson et Spencer (1966).

L'aspect fonctionnel et appliqué de l'habituation a aussi été modélisé pour des robots, reproduisant le phénomène dans ses grandes lignes, tel qu'observé expérimentalement avec les agents biologiques (Marsland *et al.* 1999). L'habituation a été démontrée dans quelques contextes seulement, dont l'ajustement de mouvements latéraux lors de déplacements de robots dans des corridors étroits (Chang, 2005). Incidemment, l'ajout d'une fonction d'habituation a permis une navigation plus fluide, en atténuant les stimuli perçus par les senseurs de distance. Normalement sans ce type d'apprentissage, le mode réactif aurait conduit les robots à des changements brusques d'orientation.

L'habituation a aussi été reliée au phénomène de l'attention par un processus biomimétique appliqué à la vision des robots (Marsland *et al.* 1999). Dans cette expérimentation, des robots s'habituèrent avec la fréquence des stimuli visuels conformément à une hiérarchie de mémoires, distribuée selon l'intensité (biais neuronal) et la localisation des stimuli (plusieurs cellules photo-électriques). La réponse comportementale à la nouveauté était de se tourner vers la lumière par phototaxie, pour graduellement ignorer les stimuli déjà rencontrés et se concentrer sur les nouveautés. Dans cette même perspective, l'habituation a aussi été implémentée comme modèle de détection de nouveautés pour des robots (Marsland *et al.* 2005) et des interfaces robots-utilisateurs (Déniz *et al.* 2004).

C'est dans cette logique d'inclure l'habituation dans des RNAI agissants comme contrôleur de robots qu'un des articles de la thèse a été développé, comme premier outil d'apprentissage en robotique bio-inspirée. Tel que mentionné ci-haut dans les quelques



exemples d'applications pour des robots, ceux-ci devraient trouver avantage à posséder cette fonction d'apprentissage (Marsland, 2009), notamment pour éviter la computation inutile de stimuli répétitifs, en rehaussant ceux qui sont nouveaux et en exploitant au maximum les diverses propriétés temporelles des informations perçues. Ainsi, l'habituation devient une fonction d'apprentissage adaptative comportementale généralisable qui devrait être autant valide pour le monde naturel qu'artificiel.

Les motivations de proposer un nouveau modèle d'habituation ont d'une part, été celles d'adapter cette fonction naturelle dans une perspective spécifique d'utilisation dans des RNAI et d'autre part, d'une optique d'implémentation de la fonction dans un corpus d'apprentissage de base pour des robots. À la lecture des articles du domaine, il s'est avéré que certaines lacunes des modèles computationnels actuels existaient, conduisant à la pertinence d'introduire des aspects novateurs. Ainsi, le modèle d'habituation proposé dans cette thèse permet d'élargir le spectre des motifs temporels reliés à certaines variables du processus, en regard des données expérimentales. À notre connaissance, aucun modèle computationnel de l'habituation pour des RNAI n'était encore capable de traiter autant de motifs temporels hétérogènes de stimuli, soit le mode qui est le plus fréquemment rencontré dans le monde réel.

Dans un premier temps, il a d'abord fallu reproduire et valider les propriétés de base de l'habituation à travers des algorithmes de programmation et ensuite les intégrer à un RNAI. La conjonction des thèmes de recherche, habituation et réseaux de neurones, a déjà été explorée pour des RNA traditionnels (Stiles et Ghosh, 1997; Chang, 2005), mais très peu de chercheurs ont exploité l'habituation en lien avec les RNAI et leur capacité computationnelle intrinsèque augmentée dans le domaine de la temporalité. Aussi, plus rares sont les chercheurs qui ont tenté d'incarner un modèle d'habituation pour des robots et encore moins ont opté pour le paradigme de la triple combinaison habituation-RNAI-robots, laissant un vide à combler dans ce champ spécifique de l'apprentissage en robotique bio-inspirée.

Tout comme sa contrepartie biologique, le modèle computationnel d'habituation proposé dans l'article a été réalisé simultanément sur plusieurs niveaux conceptuels.

Notamment, le mécanisme de décrémentation de la réponse ainsi que l'orientation vers le type de mémoires sont basés sur des catégories temporelles implicites de l'IIS. Les effets de la fonction introduite sont observés du niveau de changements de l'efficacité synaptique au niveau comportemental des robots. La conclusion est que le modèle est valide et englobe tous les types de patrons temporels d'IIS, avec une réponse comportementale complètement bio-inspirée, similaire aux données expérimentales.

### 2.3 Thématiques connexes à explorer

L'habituation est une fonction d'apprentissage primaire représentée par une diminution de la réponse comportementale lorsque des stimuli non vitaux sont répétés. Le mécanisme interne de ce phénomène repose sur la dépression synaptique et plus spécifiquement, sur la diminution des éléments constitutifs et fonctionnels présynaptiques, en lien avec la fréquence des IIS. Ce processus diffère de la sensibilisation qui lui repose avant tout sur une facilitation synaptique en regard de la répétition de stimuli à fortes amplitudes ou nociceptifs (Antonov *et al.* 2010).

Cet autre phénomène d'apprentissage non associatif primaire devrait être aussi modélisé. Dans le monde naturel, les stimuli entrants sont de types variés et fréquences multiples, conduisant à un chevauchement permanent de l'habituation et de la sensibilisation. Dans la perspective de réaliser une IA bio-inspirée complète, l'inclusion du phénomène de sensibilisation ainsi que la déshabituation (Rankin et Carew, 1988) dans le modèle computationnel actuel pourrait permettre de discerner les chevauchements de ces formes d'apprentissage et même aider à comprendre l'interaction de ces fonctions dans le monde biologique. Trouver des applications concrètes en robotique et justifier cet ajout de la sensibilisation au modèle sont des questions qui demeurent toutefois à être approfondies, car moins évidentes que pour l'habituation. Ce qui est également reflété par une absence de littérature sur l'intégration d'un modèle de la sensibilisation à des modèles actuels sur l'habituation dans le domaine de la robotique.

Dans un autre ordre d'idée, il serait intéressant de comprendre les mécanismes biologiques qui guident les changements au niveau moléculaire, entre l'augmentation ou la diminution de la réponse comportementale en regard de la nature et de l'intensité d'un stimulus ? Sur quel critère un stimulus est-il spécifique et n'est plus considéré similaire ? À cet égard, des recherches sur l'habituation des odeurs chez la souris (McNamara *et al.* 2008) indiquent que la structure des odeurs discrimine la spécificité du stimulus. D'autre part, est-ce qu'il existe une valeur cible d'intensité absolue ou relative qui enclenche soit la sensibilisation ou l'habituation ? Comment ces deux processus interagissent-ils ?

Aussi, un aspect encore peu connu de l'habituation qui mériterait une clarification, est celui du passage de la mémoire à court terme vers le long-terme. Bien qu'une piste reliant les IIS et les expérimentations sur des blocs d'apprentissages offre une corrélation phénoménologique, l'explication au niveau moléculaire manque actuellement. Il doit exister des valeurs seuils qui aiguillent la bifurcation d'un type de mémoire à l'autre. Un suivi de ces paramètres dans la littérature aiderait éventuellement à réaliser un modèle computationnel plus complet. Il en va de même pour le phénomène de reconsolidation de la mémoire à long terme, lorsqu'il y a sollicitation par de nouveaux stimuli (Cai *et al.* 2012).

Également, la plupart des expérimentations sur l'habituation ont été réalisées sur des circuits neuronaux réflexes sensori-moteurs, donc de premier ordre (Simons-Weidenmaier *et al.* 2006). Très peu d'études (Engel et Wu, 2009; Larkin *et al.* 2010) relatent le phénomène au niveau de synapses centrales situées entre les neurones sensoriels et des interneurones ou entre interneurones. Il en est de même sur la contribution de l'habituation aux différents réseaux parallèles aboutissant à un même motoneurone. Est-ce que l'habituation est un mécanisme propre au type particulier des cellules sensorielles ? Est-ce que l'habituation s'exprime dans tous les types de cellules neuronales ou seulement sur certaines cellules spécifiques ? Si l'habituation est ubiquitaire, il en résulterait que ce mécanisme agit comme régulateur intrinsèque de la fréquence des PA en les atténuant aussi selon leur IIS. Si cette hypothèse s'avère être

vraie, il s'agirait d'une propriété neuronale qui aurait certainement de grandes répercussions au niveau des circuits et des comportements, recadrant notre compréhension actuelle du processus de l'information. Il va s'en dire qu'au niveau computationnel, les avantages devraient être les mêmes à incorporer cette propriété neuronale théorique dans des RNAI et pour des robots bio-inspirés.

Aussi, le rôle modulateur extrinsèque des neurones inhibiteurs dans la décrémentation des comportements est observé et impliqué dans certains circuits, parfois en ajout du processus d'habituation (Brystol et Carew, 2005), parfois sans aucune trace de ce type d'apprentissage (Agin *et al.* 2006). Ainsi, il est possible d'inhiber un comportement par simple association passive d'évitement. La diminution comportementale, lorsqu'elle est l'objectif à simuler, se résout au moins par trois mécanismes différents, soit l'adaptation sensorielle, l'habituation et les circuits impliquant des neurones inhibiteurs. Une réponse comportementale diminuée le moins complexe résulte probablement de la superposition de ces différents mécanismes internes évoluant sur des échelles temporelles et cinétiques différentes. Il serait pertinent de capitaliser sur cette diversité phénoménale biologique, dans la réalisation d'un modèle computationnel complet sur l'apprentissage en robotique bio-inspirée, particulièrement pour l'implémentation sur les plates-formes physiques.

Hormis les conditions de laboratoire, l'habituation n'est pas un processus isolé. Les neurones sont des éléments dynamiques et un phénomène de maintenance module les changements d'excitations membranaires en permanence, même observés lors d'habituation (Fischer *et al.* 2014). En fait, tous les types d'apprentissage dans toutes les modalités sensorielles se déroulent en concomitance, et ce, sur une gamme d'échelles temporelles étendues (Szyszka *et al.* 2008; Çevik, 2014).

Dans un but de mimer le réalisme biologique, l'intégration de plusieurs types d'apprentissage est une avenue de recherche peu explorée encore (Balkanius, 2000; Lloyd *et al.* 2014). Le domaine de la robotique bio-inspirée pourrait favoriser ce genre d'étude intégrative des fonctions d'apprentissage de par son mode analytique privilégié et l'accessibilité des données produites. C'est dans cette perspective d'intégration des



phénomènes d'apprentissage en robotique bio-inspirée que le modèle d'habituation de cette thèse a été développé. Ainsi, au modèle global d'apprentissage par modulation synaptique, pourrait s'ajouter d'autres fonctions telles que l'adaptation sensorielle, la sensibilisation, la déshabituatation, le conditionnement classique, le conditionnement opérant et potentiellement aussi des types d'apprentissage de plus haut niveau.

Cette intégration permettrait aussi d'étudier les interrelations des différents mécanismes d'apprentissages. Entre autres, il a été démontré que si un stimulus neutre est préalablement habitué et postérieurement associé à un stimulus inconditionnel dans un protocole de CC, alors la force consécutive des nouveaux liens créés conduira à une réponse comportementale qui est diminuée (Çevik, 2014). Ces observations sont également valables pour la sensibilisation et le conditionnement opérant. Cet effet de jumelage des processus n'est pas strictement additif. Aussi, tous les mécanismes d'apprentissages dépendent des états internes, systèmes de valeurs ou motivations des agents biologiques. Ainsi, selon le contexte, un même protocole d'habituation, de CC ou de CO, aboutira à des observations comportementales différentes selon les états internes. Ces états internes sont nécessairement reliés à d'autres neurones qui influencent et modulent les changements synaptiques. Ainsi, les apprentissages de plus haut niveau tel que le CC et le CO sont influencés par l'historique des formes d'apprentissages de plus bas niveau hiérarchique tel que l'habituation et la sensibilisation, tout comme les contextes de valeurs internes. Or, aucune étude n'est actuellement disponible sur les interdépendances entre les différents mécanismes d'apprentissage et leurs impacts au niveau comportemental autant au niveau biologique que robotique. C'est précisément l'axe principal exploré dans la thèse, que de développer un ensemble compatible de fonctions d'apprentissage en bio-inspiration robotique afin d'étudier ultérieurement les possibles interrelations et leurs impacts.

Un autre sujet d'intérêt en lien avec l'habituation est celui du bruit ambiant. Des chercheurs ont démontré qu'il y avait peu d'habituation lorsque les stimuli sont chaotiques, multiples et dont l'occurrence s'établit sur plusieurs échelles temporelles. Pour des robots évoluant dans le monde réel, il semble primordial d'avoir un modèle

d'habituation cohérent avec les données biologiques observées empiriquement, car peu importe la modalité sensorielle en cause, cette variable du bruit ambiant est indéniablement présente dans la nature.

Aussi, l'habituation est reconnue comme ayant une capacité indirecte et intrinsèque de détection de la nouveauté par atténuation des stimuli persistants. Il pourrait s'agir d'une étape préliminaire permettant d'orienter de façon réflexe, l'attention vers un objet plutôt qu'un autre. Ainsi, ce mécanisme permettrait potentiellement d'influencer la sélection d'actions et même la prise de décisions. À cet égard, un article complémentaire, mais en lien avec la thèse a été produit sur ce sujet (Voir annexe E).

Dans cette perspective, bien qu'il soit intéressant de vérifier ces hypothèses dans des conditions expérimentales contrôlées se déroulant sur des stimuli unimodes, il serait encore plus pertinent de pousser le raisonnement sur des robots possédant un répertoire de plusieurs actions hiérarchisées lorsque mis en contexte avec des motifs de stimuli multisensoriels en provenance d'objets complexes.

Enfin, comme les échelles temporelles diffèrent grandement d'un protocole expérimental à l'autre concernant la durée des stimuli, le nombre de stimuli pour obtenir un maximum d'habituation, les IIS, le temps de récupération, les durées des mémoires à court et long terme, la modalité sensorielle, l'espèce et le type d'animal étudié, alors comment et sur quelle base devrait-on paramétrer un modèle computationnel de l'habituation chez un robot ? Comment généraliser la valeur d'un signal lorsque transformé en un motif de PA ? Autant de questions ouvertes qui trouveront probablement une réponse dans la contrepartie naturelle...



## ARTICLE 1

Habituation: a non-associative learning rule design for spiking neurons and an autonomous mobile robots implementation

Bioinspiration & biomimetics

## Résumé

Cet article propose une nouvelle fonction d'habituation applicable aux robots contrôlés par des réseaux de neurones artificiels à impulsions. Cette règle d'apprentissage de type non associatif est modélisée au niveau synaptique. Elle est validée par les modifications de comportements de robots physiques et virtuels dans un monde 3D, consécutivement à l'application de différents motifs de stimuli. L'habituation est minimalement représentée par une réponse atténuée après une exposition répétée de stimuli externes persistants. Basée sur les récentes données de la littérature en neuroscience, l'originalité de cette règle d'apprentissage repose sur une réponse qui est modulée en fonction de différentes fréquences d'application de motifs de stimuli. Le filtrage de stimuli répétitifs par le phénomène naturel de l'habituation a été suggéré comme étant un facteur clé dans le phénomène de l'attention. Insérer une telle règle d'apprentissage fonctionnant sur de multiples échelles temporelles de stimuli pourrait ainsi, en augmenter d'autant les capacités adaptatives comportementales des robots, en ignorant un plus large éventail d'informations contextuelles qui ne sont pas pertinentes.

(Voir l'annexe B pour l'article intégral)

## CHAPITRE III

### CONDITIONNEMENT CLASSIQUE

#### 3.1 Volet psychologie

Une autre forme élémentaire d'apprentissage est celle du conditionnement classique (CC). Cette procédure issue du domaine de la psychologie implique une association entre un stimulus initialement neutre (*neutral stimulus*) (NS) provenant généralement de l'environnement et un stimulus inconditionnel (*unconditioned stimulus*) (US) qui est essentiel pour l'organisme. La nature de l'US peut-être de type aversif ou récompense et induit une réponse inconditionnelle naturelle (*unconditioned response*) (UR), souvent de forme réflexe. Une des théories expliquant le CC est celle que le CS se substitue à l'US, l'UR agissant comme renforçateur dans l'association entre le NS et l'US. Pavlov (1927) a fait figure de pionnier dans ce type d'apprentissage en exploitant le comportement du chien qui salive (UR) devant de la nourriture (US).

L'exemple méthodologique canonique a été de faire sonner une cloche à diverses reprises juste avant de nourrir un chien. Pavlov avait alors démontré que suite à cette procédure d'apprentissage par conditionnement et un certain temps d'acquisition, le seul son de la cloche (*conditioned stimulus*) (CS), sans être suivi de nourriture (US), suffisait à obtenir une réponse conditionnée (*conditioned response*) (CR), soit de faire saliver le chien. L'interprétation résultante est que le chien aurait passivement appris que le son d'une cloche devenait prédictif de nourriture, anticipant un renforcement positif dans ce contexte. Les autres recherches de Pavlov l'ont conduit à élaborer une théorie et à suggérer des mécanismes biologiques du CC, avec les diverses relations temporelles et de magnitudes qui existent entre les CS, US, CR et UR, donnant :

US → UR      NS + US → UR      CS → CR

Le modèle de CC a été raffiné par les psychologues contemporains tels que Rescorla et Wagner (1972) en incluant plusieurs phénomènes secondaires. Cependant, le CC demeure mal compris, notamment pour certains phénomènes connexes du domaine observés expérimentalement (Schmajuk, 2010). Entre autre, le phénomène de généralisation implique qu'une CR peut être obtenue également par des CS partageant des caractéristiques semblables au stimulus neutre de base. Aussi, il semble possible d'obtenir sous certaines conditions, une discrimination de stimulus par CC. Une autre caractéristique importante reliée au CC est celle de l'extinction, c'est-à-dire qu'une CR diminue ou disparaît si le CS est présenté souvent sans être suivi du US. Dans cet ordre d'idée, après une période d'extinction définie, une récupération spontanée de la CR peut être observée sous certaines conditions. Bien que chacune de ces observations empiriques exige un redressement du modèle théorique du CC, l'élucidation des mécanismes sous-jacents à l'échelle cellulaire est encore loin d'être accompli.

### 3.2 Volet neuroscience

Le CC implique l'établissement d'un lien associatif entre un stimulus neutre et une réponse conditionnée quasi identique à celle de la réponse inconditionnelle obtenue lors de la présentation d'un stimulus inconditionnel. En général, cette association est observée lorsqu'il y a répétition de l'association entre le NS et l'US. Il existe toutefois des cas où l'apprentissage peut être réalisé lors d'une unique phase d'acquisition (Alexander *et al.* 1984; Andrew et Savage, 2000; Michel *et al.* 2008; Sugai *et al.* 2007).

Les chercheurs tentent de trouver le substrat et les mécanismes impliqués dans la formation de ce lien associatif. Hebb (1949) a été un des premiers à formuler des hypothèses au niveau cellulaire dans le phénomène du CC, impliquant que des changements sont observables entre deux neurones s'ils sont simultanément actifs. Il a cependant fallu attendre plusieurs décennies avant de mieux comprendre les bases biochimiques et les transformations cellulaires impliquées dans le CC.

À l'aide de modèles animaux possédant des systèmes nerveux simples, tels que le ver nématode *Caenorhabditis elegans*, l'escargot *Lymnaea stagnalis*, la mouche *Drosophila melanogaster* et la limace de mer *Aplysia californica*, les hypothèses sur la nature du lien associatif pointent actuellement soit vers une ou des molécules de coïncidence des événements dans un même neurone (Brembs et Plendl, 2008).

Dans ces modèles réductionnistes, l'apprentissage par CC résulte d'une modification au niveau moléculaire et éventuellement cellulaire, témoignant de la séquence temporelle ordonnée et spécifique des événements passés. Ce lien associatif représente une forme de mémoire implicite, qui perturbe l'homéostasie neuronale (Kandel, 2001). Cette altération de causalité qui repose sur un changement structurel et consécutivement de la fonction neuronale se répercute parfois sur le réseau entier et pouvant aussi mener à l'observance d'un changement comportemental au niveau de l'agent complet.

Une caractéristique importante du phénomène de la mémoire et de l'apprentissage est celle de la plasticité synaptique, lieu présumé du changement correspondant à la trace mnésique (Mayford *et al.* 2012) et du lien associatif impliqué dans le CC. En effet, la synapse est une structure dont la diversité traduit une variabilité de fonctions. La synapse est aussi un élément neuronal fortement dynamique (Smythies, 2002), continuellement modulée par des signaux homo et hétérosynaptiques. En ce sens, l'activité neuronale présynaptique conditionne la réponse postsynaptique par plusieurs mécanismes correspondant à la plasticité synaptique (Zucker et Regehr, 2002). Parmi la variance des cinétiques d'activités neuronales, certaines ont une corrélation positive entre l'ordre temporel d'émission de PA de deux neurones liés et le changement de l'efficacité synaptique consécutif.

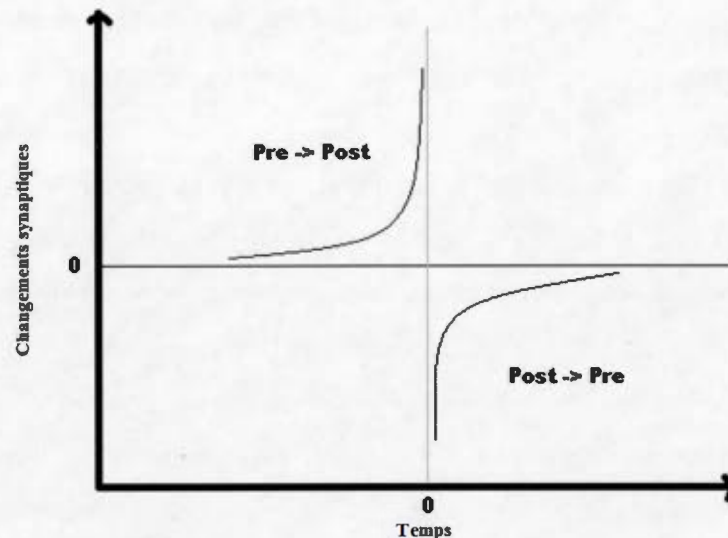
En exemple, le réflexe de retrait du siphon chez l'aplysie est plus fort et dure plus longtemps si une touche légère est précédée d'un stimulus nociceptif électrique (Antonov *et al.* 2001). Chez cet invertébré, la modulation du réflexe de base a été démontrée *in vitro* et *in vivo* dans un protocole expérimental sensori-moteur à un seul lien monosynaptique. Le changement neuronal consécutif de sensibilisation a été

constaté au niveau synaptique avec des mesures observables et quantifiables du changement, se répercutant au niveau comportemental de l'animal. La durée de la persistance de cette modification synaptique a aussi été démontrée comme étant sensible aux patrons spatio-temporels des stimuli entrants ainsi qu'aux activités neuronales sous-jacentes, conduisant à la description d'une hiérarchie de mémoires (*short- intermediate- long-term memory*) (Hawkins *et al.* 2006).

À plus petite échelle, les diverses cascades biochimiques intracellulaires observées lors de l'apprentissage par CC impliquent des échanges de communications avec le matériel génétique intranucléaire qui sont responsable des modifications synaptiques plus permanentes (Kandel, 2001). Ces modifications synaptiques reposent sur la synthèse de nouvelles protéines, conduisant à altérer l'efficacité d'une synapse et/ou en modifier le nombre. Cette modulation persistante dans le temps de l'efficacité synaptique rend plus facilement observable le phénomène de CC lors des expérimentations.

Une des hypothèses soutenant le CC comme mécanisme modifiant l'efficacité synaptique est celle du STDP (Caporale et Dan, 2008; Markram *et al.* 2011), un modèle cellulaire théorique robuste ayant des propriétés associatives qui corroborent les intuitions de Hebb. En effet, si une cellule neuronale émet un PA vers une autre dans une courte fenêtre temporelle déterminée (plus ou moins 50 msec), alors il y aura modification des synapses unissant ces deux neurones (Figure 3.1). L'ordre temporel de la séquence des PA détermine le sens de la modification synaptique, soit une potentialisation pour une paire pré-post de PA et une dépression synaptique pour une paire post-pré. Plus le temps entre les deux PA est court, plus la modification sera forte et inversement. Ainsi, le modèle de STDP vérifie de façon raisonnable, un ensemble de données obtenues par des expérimentations effectuées sur plusieurs types de cellules *in vitro* et *in vivo* (Usrey *et al.* 1998), entre autres sur des neurones corticaux pyramidaux de la couche V et sur des neurones de la région CA1 de l'hippocampe (Abbott et Nelson, 2000; Markram *et al.* 1997; Shouval *et al.* 2002; Sjöström *et al.* 2002).





**Figure 3.1** STDP. Courbes qualitativement représentatives des changements synaptiques observés lors de l'application d'un protocole de stimulation électrique, temporellement et séquentiellement associée dans des neurones pré-post synaptiques (Schématisation adaptée d'après Bi et Poo, 1998).

Dans un modèle STDP, la coïncidence entre des paires de PA provenant de neurones pré-post ou post-pré synaptiques, suppose une « rencontre » biochimique dont la nature est présumée être différents taux de concentrations moléculaires avec comme résultante, un changement local de l'efficacité synaptique. Ce bref changement d'efficacité de la synapse activité-dépendante est produit par des modifications fonctionnelles et structurelles. Entre autres dans certains neurones, cette modulation reflète une substitution des sous-types et du nombre relatif de canaux AMPA/NMDA ainsi qu'un changement morphologique du bouton synaptique et même de l'arbre dendritique, ce qui modifie la cinétique éventuelle des PPSE/PPSI (mémoire à court terme) reçus par le neurone postsynaptique.

Par exemple, certains neurones émettent des neurotransmetteurs excitateurs de type glutamate consécutivement à l'émission d'un PA. Ces molécules possèdent une affinité pour des récepteurs ioniques de type NMDA situés sur l'élément postsynaptique. Ces canaux perméables aux cations ( $\text{Na}^+$ ,  $\text{K}^+$  et  $\text{Ca}^{++}$ ) sont sensibles au voltage et

normalement bloqués par la présence physiologique extracellulaire d'ions  $Mg^{++}$ . Lorsqu'il y a une dépolarisation neuronale postsynaptique, les ions  $Mg^{++}$  sont délogés, laissant les pores membranaires des canaux NMDA ouverts et permettant une entrée de  $Ca^{++}$  (suivie d'une autre par l'entremise de canaux  $Ca^{++}$  voltage-dépendant). Ce changement de voltage postsynaptique requiert non seulement un PA postsynaptique, mais aussi l'aide des canaux AMPA également sensibles au glutamate. Les canaux AMPA possèdent une cinétique de PPSE plus courte que les canaux NMDA, mais de forte intensité, servant à amplifier le signal présynaptique. L'arrivée de glutamate présynaptique et la dépolarisation postsynaptique trouvent une résonance comme détecteur associatif de coïncidences. Dans des étapes successives, des cascades biochimiques plus complexes impliquent des niveaux de concentration de  $Ca^{+}$  et une rétropropagation du signal postsynaptique. Ces messages ciblés localement aux synapses impliquées causent le changement durable et observable de l'efficacité synaptique soit vers une potentialisation ou une dépression (Rudy, 2014).

Ce mécanisme de ciblage synaptique spécifique est encore actuellement au stade d'hypothèse de recherche avec des alternatives au modèle théorique de base (Izhikevich, 2007; Redondo et Morris, 2011; Pöpper *et al.* 2011). L'idée générale est que certaines synapses subiraient des modifications temporaires leur conférant des propriétés mnésiques à court terme. Puis, lorsqu'identifiées dans un contexte précis et stimulées à nouveau, elles entreraient dans une phase de consolidation et même de reconsolidation pour des passages plus durables vers de la mémoire à long terme. Une meilleure connaissance de tous les états synaptiques possibles en corrélation avec leurs différents potentiels temporels mnésiques, ainsi que les mécanismes opérant chacune des transitions, représentent des clés maîtresses du problème.

La persistance du changement de l'efficacité synaptique trouve aussi un corollaire avec le phénomène de LTP (fréquence de stimulation 100-200 Hz) et *long-term depression* (LTD) (1-5 Hz), effet STDP supposé de prolongement qui est observé dans le temps (Bliss et Lømo, 1973). Plusieurs phases temporelles ont d'ailleurs été constatées avec LTP/LTD (Lüscher et Malenka, 2012), impliquant d'abord des

changements synaptiques transitoires mineurs (phase précoce), puis une participation du matériel génétique local (phase intermédiaire), et finalement central (phase tardive). Ces changements synaptiques évoluant sur différentes échelles de temps semblent être impliqués dans la trace mnésique et représenteraient la base du mécanisme de l'apprentissage (Rudy, 2014). Plusieurs types d'expérimentations (Dan et Poo, 2004) ont d'ailleurs démontré un lien entre les phénomènes STDP, NMDA dépendant, LTP et l'observation de changements comportementaux significatifs.

Outre le pairing de PA et l'effet de fréquence des PA selon les types de neurones, la composition structurelle et la localisation synaptique, les neurotransmetteurs impliqués (c.-à-d. Glutamate, Gaba) et les patrons en volée des PA, découlent une grande variété de sous-types de fonctions STDP qui reflète un riche ensemble de mécanismes internes répondant aux diverses communications neuronales possibles (Caporale et Dan, 2008). De ces variantes de modèles STDP résulte l'observation de différentes courbes de changement de l'efficacité synaptique, tant au niveau de l'étendue des fenêtres temporelles que dans leur magnitude (Dan et Poo, 2006). En général, un ordre de grandeur de 100 msec comme fenêtre temporelle, un changement d'efficacité synaptique de 100% (courant excitateur de 0.4 nA ou PPSE de 4mV) et une fréquence d'application de stimuli d'environ 60Hz, représentent des valeurs limites courantes en termes de données expérimentales recueillies (Lüscher et Malenka, 2012). Les paramètres et résultats recueillis semblent aussi être asymétriques, avec une fenêtre temporelle plus longue pour post-pré et une amplitude de changement plus forte avec pré-post. De plus, puisque l'activité de neurones liés affecte les synapses par des règles STDP, il en découle un aspect dynamique de modification de ces règles avec des résultats difficiles à reproduire en termes de circuits *in vivo* (Jacob *et al.* 2007). Il existe même des cas où les résultats STDP sont inversés, impliquant que des paires de PA pré-post occasionnent une baisse de l'efficacité synaptique (Bell *et al.* 1997) !

Au niveau mathématique, la règle de base modélisant la phénoménologie du pairing temporelle de PA dans le STDP s'exprime par une fonction exponentielle double :

$$W(x) = \pm A * e^{(x/\tau)}$$

où,

- W = variation du poids synaptique
- x = intervalle de temps entre les PA pré-post ou post-pré
- A = sens du changement synaptique (+ pour pré-post/ – pour post-pré)
- $\tau$  = constante de l'ordre de 10 msec reflétant les données expérimentales

À cette règle de base du phénomène STDP s'ajoutent aussi des variantes qui reposent sur d'autres prémisses que le simple pairage unitaire de PA pré-post ou post-pré (Feldman, 2012). En effet, il a été démontré que lors de certains protocoles conduisant au phénomène de LTP/LTD, l'application de paires consécutives de PA à fréquence définie, offre des caractéristiques non seulement de changement de l'efficacité synaptique en regard d'une séquence et d'un ordre temporel précis de cette paire de PA, mais aussi en regard à la fréquence et composition interne du patron de ces PA lorsqu'il y en a plusieurs. Ainsi, des triplets de PA (pré-pré-post, pré-post-pré, post-pré-post ou post-post-pré) représentent d'autres possibilités de communication neuronale et conduisent aussi à des changements de l'efficacité synaptique, qui diffèrent du simple pairage dans la cinétique observée. De plus, dans le cas où la fréquence de PA est stable, il a été démontré que les premières paires de PA ou triplets ont plus d'effet sur le changement synaptique que les suivantes avec, encore une fois, une relation de type fonction non linéaire décroissante (Froemke *et al.* 2005). Dans le cas pré-post-pré, aucun changement synaptique n'est constaté. Cependant, dans le cas post-pré-post, une potentialisation de l'efficacité synaptique est observée (Froemke et Dan, 2002), ce qui implique un gain de la potentialisation sur la dépression synaptique. Aussi, l'apprentissage de masse versus l'apprentissage par protocole espacé, influence grandement les variations de poids synaptiques; ce qui exige somme toute, un modèle STDP qui tient compte de toutes ces variables, dont la fréquence des paires de PA ainsi que de l'effet du taux possible de récupération de ces synapses, vers leurs états potentiels homéostatiques.

Ainsi, un ensemble de données expérimentales (Bi et Poo, 1998; Sjöström *et al.* 2001, Wang *et al.* 2005) remet en cause la règle simple des paires de PA parce que minimalement, la prédiction du changement de l'efficacité synaptique obtenu par des protocoles de triplets ou quadruplets de PA, est erronée. Premièrement, une fréquence faible de PA avec délai ne peut prédire le phénomène LTP avec une formule STDP standard. Deuxièmement, si plusieurs paires de PA sont rapprochées dans le temps, l'interaction avec les suivantes occasionne au niveau computationnel, l'annulation des effets sur la synapse. Troisièmement, que ce soit par des paires de PA, de triplets ou de quadruplets, une asymétrie des résultats se constate expérimentalement, mais ne peut se vérifier computationnellement par une formule générale de STDP. En ce sens, des modèles théoriques de STDP plus complexes ont été élaborés (Pfister et Gerstner, 2006; Gjorgjieva *et al.* 2011) avec plus ou moins de succès.

À cet égard, en addition au changement de l'efficacité synaptique selon une dépendance temporelle entre des PA pré-post ou post-présynaptique, la modélisation computationnelle d'une dépendance de fréquence individuelle de PA pré et post ainsi qu'un effet de fréquence des paires de PA représentent l'essentiel de la contribution de l'article sur STDP dans le CC (voir plus bas). Cette règle STDP élargie de type interaction totale démontre une capacité à gérer tous les types de motifs de PA par l'inclusion de plusieurs règles non linéaires imbriquées. De plus, cette règle a été vérifiée par des applications concrètes d'IA complètes virtuelles et physiques. Il est à noter que, pour cette règle STDP du modèle présenté dans cette thèse, aucune référence à des paramètres biophysiques n'est explicite, favorisant plutôt la phénoménologie et les fonctions ainsi que la cohérence temporelle entre les différentes variables.

Bien qu'encore incomplète en termes de détails bio-inspirés, cette fonction STDP permet d'être en accord avec un plus grand nombre de données empiriques. Cette règle d'apprentissage est notamment mieux adaptée aux signaux complexes issus du monde réel comme ceux retrouvés lors d'implémentations sur des robots physiques. Outre la vérification du postulat de Hebb avec une coïncidence dans le temps de PA, des cas de LTP ont aussi été rapportés uniquement avec une stimulation présynaptique (Bliss et



Collingridge, 1993), ce qui vraisemblablement requiert l'inclusion d'une règle de changement de l'efficacité synaptique qui est aussi indépendante du pairage de PA. En l'occurrence, le besoin d'un modèle d'apprentissage qui inclue une règle associative et non associative.

Donc, à ces différentes dépendances temporelles associatives de PA par des règles STDP, le CC doit être compris et étudié dans un modèle d'apprentissage plus large, incluant aussi les phénomènes non associatifs résultant de la plasticité présynaptique telle que l'habituation et la sensibilisation, conduisant respectivement à une dépression et une facilitation synaptique par répétition de stimuli entrants, de faibles et fortes amplitudes. Pour cette thèse, lorsque la règle STDP comme mécanisme du CC est conjuguée avec les changements présynaptiques de l'habituation précédemment discutée, l'adaptation des robots dans le domaine temporel est grandement accrue dans le modèle global de fonctions d'apprentissage bio-inspirée.

### 3.3 Volet IA

L'IA comme discipline scientifique vise à simuler l'intelligence naturelle. Dans le paradigme de la bio-inspiration et plus spécifiquement dans le champ de la neuro-robotique, la simulation du CC est basée sur les informations actuelles en neurosciences (Dan et Poo, 2006; Caporale et Dan, 2008) et les modèles computationnels disponibles (Balkanius et Morén, 1998; Schmajuk, 2010). Seulement quelques recherches ont été faites sur le CC en conjugaison avec les RNAI (Liu et Shapiro, 2008) ou en croisement avec le domaine de la robotique (Di Paolo *et al.* 2002; Di Paolo, 2003; Alnajjar et Murase, 2006; Alnajjar et Murase 2008; Alnajjar *et al.* 2009, Aréna *et al.* 2009).

Ainsi, basés sur les phénomènes STDP / LTP-LTD, différents modèles computationnels du CC peuvent être testés et implémentés sur des plates-formes physiques telles les robots. Il devient alors possible de vérifier les hypothèses biologiques allant d'un changement d'efficacité synaptique jusqu'à l'observation d'un changement durable du comportement pour une perspective d'agent artificiel robotisé.

C'est dans ce cadre expérimental contrôlé qu'il est permis d'aborder des motifs de PA plus complexes et d'en étudier les résultats au niveau comportemental. L'analyse de la trace d'une modification synaptique dans un circuit neuronal biologique complexe à partir des entrées sensorielles jusqu'aux actions motrices est actuellement impossible chez un animal, mais tout à fait possible pour un robot!

Récemment, des chercheurs (Soltoggio *et al.* 2013) ont proposé un mécanisme alternatif à STDP, soit des corrélations statistiquement rares, pouvant expliquer des associations entre une récompense, des stimuli et des actions. Ce processus également basé sur la plasticité synaptique implique une gestion des différentes traces des synapses éligibles en regard des bons éléments à lier. Cette étude démontre du CC et du CO avec un robot physique (iCub) supervisé par un humain, offrant un contexte robuste en regard des différents délais impliqués dans l'application des stimuli, des actions et des récompenses.

En l'instar du sujet de cette thèse, il est intéressant de constater que d'autres chercheurs tentent aussi d'intégrer différents mécanismes d'apprentissage dans un même paradigme de robotique bio-inspirée et de RNAI. La similitude des tâches accomplies en temps réel ainsi que la gestion des événements sur différentes échelles temporelles entre les articles de cette thèse et celui-ci suggère aussi que le niveau de complexité actuel du domaine est encore minimaliste. Cependant, cet article n'incluait pas de fonctions d'habituation et les réseaux de neurones pour le CC et le CO étaient séparés, principalement pour une raison de différence des données entrantes.

En résumé pour cette section, la contribution de l'article sur le CC à cette thèse a été de conceptualiser un modèle basé sur un mécanisme biologiquement inspiré de modulation synaptique de type STDP. Cette règle d'apprentissage STDP inclue la gestion de paires individuelles pré-post de PA, de l'effet individuel pré et post des PA ainsi que l'effet cumulatif de multiples paires de PA. Cette règle d'apprentissage adaptée pour des RNAI a été validée pour des robots, de la synapse au comportement.

## ARTICLE 2

## Classical conditioning in different temporal constraints: an STDP learning rule for robots controlled by spiking neural networks

## Adaptive behavior

## Résumé

Cet article propose d'investiguer certains comportements adaptatifs d'agents cognitifs de type robot physique et virtuel. Le contexte de l'étude porte sur l'association temporelle entre des stimuli en provenance d'indices sensoriels et ceux d'actions réflexes simples. Cette association est modélisée par une règle d'apprentissage novatrice de type STDP. Cette fonction est insérée dans des RNAI servant de cerveaux-contrôleurs pour les robots. Ces systèmes cognitifs bio-inspirés sont validés par différentes procédures et tâches de conditionnement classique dans un environnement virtuel 3D et où les différents paramètres reliés aux stimuli conditionnels et non conditionnels sont variés en termes de fréquences et de coïncidences temporelles. Les résultats de ces simulations sont analysés sur différents niveaux conceptuels, de la capture des stimuli sensoriels, à la génération des potentiels d'action, jusqu'à l'observation du changement de comportements des robots virtuels et physiques. Par cette règle d'apprentissage associatif, la conclusion offerte est que les robots possèdent des comportements adaptatifs plus étendus dans le domaine de la temporalité.

(Voir l'annexe C pour l'article intégral)

## CHAPITRE IV

### CONDITIONNEMENT OPÉRANT

#### 4.1 Volet psychologie

Le conditionnement opérant (CO) ou conditionnement instrumental est une forme d'apprentissage de type associatif, liant une action et sa conséquence. Il s'agit d'un concept en psychologie, issu du courant de pensée comportementaliste, initialement développé par des pionniers comme Thorndike (Thorndike, 1901) et Skinner (Skinner, 1938). Le CO a d'abord été observé expérimentalement au niveau comportemental et se caractérise par l'inclusion de divers sous-éléments conceptuels.

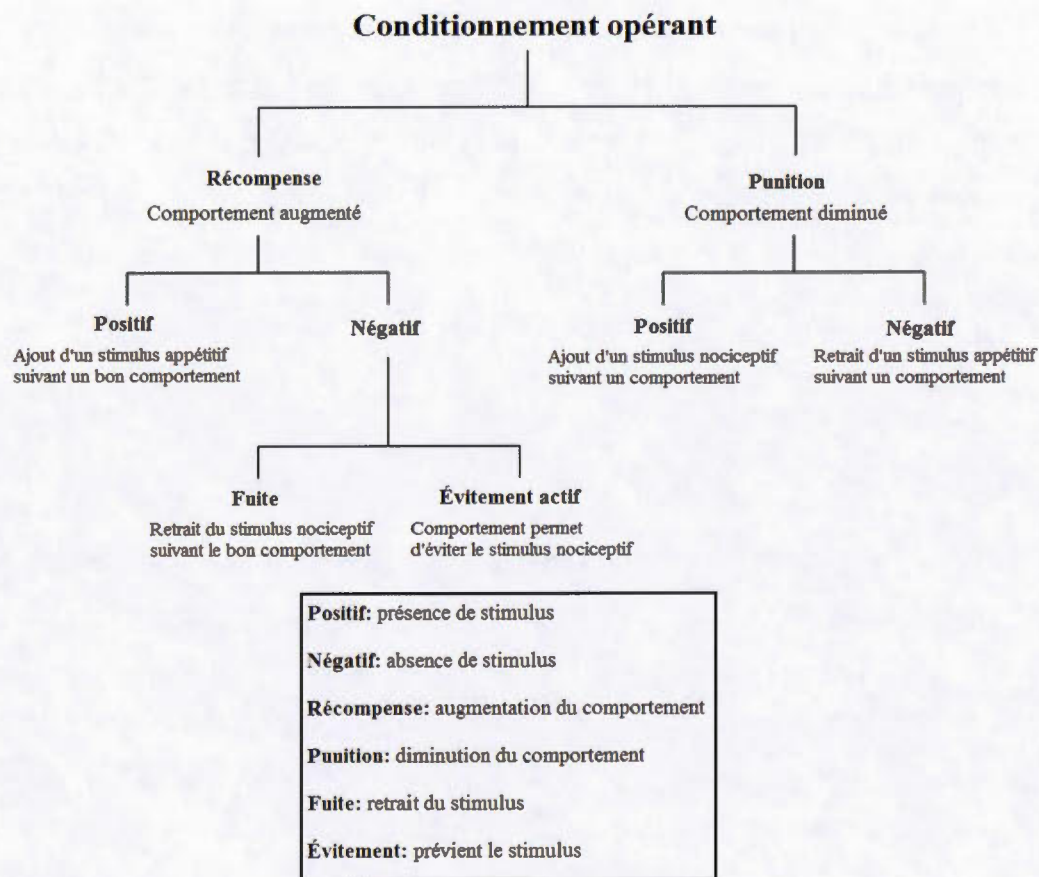
Dans une procédure de CO, la composante centrale est l'action. Cette action est accomplie par un agent cognitif dont le choix d'action est libre parmi un répertoire disponible. À l'aide de renforcements et de punitions, une action contextualisée devient privilégiée, favorisant l'acquisition d'un comportement spécifique parmi un ensemble. Il s'agit d'un objectif principal du CO que d'amener un agent cognitif à accomplir un comportement particulier, qui de façon naturelle aurait peu de chance d'être observé ou de se produire aléatoirement.

Les renforcements et les punitions sont donc le deuxième élément clé du CO. En général, un renforcement augmente la probabilité d'occurrence d'un comportement et une punition vise à en diminuer la fréquence. Il a été démontré que la force de l'association entre l'action et son renforçateur est inversement proportionnelle (fonction logarithmique) au temps de délai entre ceux-ci (Hull, 1932). Les renforcements peuvent être de type primaire s'ils sont en lien direct avec les besoins vitaux ou de type secondaire si le lien est indirect, tel qu'un objet neutre qui aurait été préalablement associé par CC à des besoins primaires.

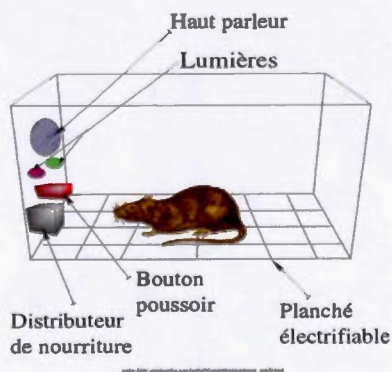
Le renforcement est positif s'il augmente la fréquence d'un comportement suite à l'ajout d'un stimulus appétitif, tandis qu'il est négatif si la fréquence d'un comportement est augmentée suite au retrait d'un stimulus aversif. La punition est dite positive si le comportement diminue avec l'ajout d'un stimulus aversif. La punition est dite négative si le comportement diminue avec le retrait d'un stimulus appétitif. La Figure 4.1 résume les quatre types possibles de procédures visant à produire du CO.

Un exemple typique de CO est illustré à la Figure 4.2. Un rat, placé dans une boîte de Skinner, munie d'un levier relié à un magasin distributeur de nourriture est encouragé à peser sur celui-ci par divers moyens de renforcement (c.-à-d. lumière, son, nourriture si le rat s'approche du levier) ou de punition (c.-à-d. choc électrique par le plancher si le rat s'éloigne du levier). Il existe énormément de variantes de ces procédures qui sont adaptées selon les espèces animales (pieuvre, abeille, pigeon, etc.) ainsi que selon les buts des expérimentateurs.





**Figure 4.1** Types de procédures dans le conditionnement opérant.  
(Tirée et traduction libre de : [https://en.wikipedia.org/wiki/Operant\\_conditioning](https://en.wikipedia.org/wiki/Operant_conditioning))



**Figure 4.2** Exemple des principales composantes utilisées expérimentalement dans le CO.  
(Tirée de : [http://fr.wikipedia.org/wiki/Conditionnement\\_opérant](http://fr.wikipedia.org/wiki/Conditionnement_opérant))

Le CO est aussi associé à plusieurs autres concepts généraux en psychologie tels que la généralisation, la discrimination, l'extinction, la récupération spontanée et le façonnement. D'un niveau plus complexe, ce dernier phénomène est observé lorsqu'un expérimentateur obtient un comportement final d'un agent à l'aide d'une série consécutive de sous-comportements construits à partir de plusieurs CO. Au final, le CO s'avère une technique utile, efficace et universelle pour faire de l'apprentissage supervisé.

#### 4.2 Volet neuroscience

Les neurosciences contemporaines tentent d'expliquer les bases biologiques du CO qui vont au-delà de l'observation comportementale, en y recherchant un substrat (Schultz *et al.* 1997). Les modèles animaux à systèmes nerveux simples d'invertébrés (Brembs, 2003) tels que la drosophile (*Drosophila melanogaster*), la limace de mer (*Aplysia californica*) et le ver nématode (*Caenorhabditis elegans*) servent présentement de phare dans ce domaine de recherche. À partir de préparations *in vitro* et *in vivo*, ces modèles animaux fournissent des explications du phénomène de CO du niveau moléculaire, génétique, cellulaire au niveau comportemental.

Les modèles de CO représentent le ou les événements constituant la coïncidence spatio-temporelle d'une action et de sa conséquence (renforcement). Bien que ces modèles animaux démontrent une capacité de faire du CO avec une résolution au niveau de petits réseaux neuronaux, certains chercheurs travaillent même au niveau de la cellule individuelle avec l'observation de liens de convergence entre les différentes cascades biochimiques (Brembs *et al.* 2002; Lorenzetti *et al.* 2008).

Par exemple, une cellule neuronale qui émet un potentiel d'action (action) et qui reçoit un influx de dopamine par la suite, modifie consécutivement l'efficacité de certaines de ses synapses ou ses propriétés membranaires (excitabilité intrinsèque), ce qui résume bien les principes du CO dans un modèle réductionniste. En général, le système de renforcement est conceptualisé au niveau cellulaire et moléculaire par des

neurotransmetteurs tels que la dopamine (Bédécarrats *et al.* 2013) et la sérotonine (Frick *et al.* 2014). Ainsi, comme autre type d'apprentissage associatif, le CO se distingue du CC au niveau conceptuel et cellulaire (Baxter et Byrne, 2006; Brembs et Plendl, 2008, Mozzachiodi, 2013).

#### 4.3 Volet IA

Le CO est aussi un phénomène d'apprentissage qui est étudié et simulé en IA. La pertinence d'émuler les processus internes naturels sous-jacents au phénomène du CO tient au fait que les modèles appliqués en IA pourraient bien souffrir de lacunes en élaguant des subtilités de la réalité biologique. En effet, certains de ces processus pourraient s'avérer être influents et déterminants.

À l'aide de RNA classiques, différentes implémentations de CO ont été développées (Sutton et Barto, 1998). D'autres variantes utilisant des RNAI ont aussi modélisé le CO. Incidemment, l'implémentation de ces modèles dans des agents artificiels complets virtuels (Fiore *et al.* 2008) et physiques a été explorée dans le domaine de la robotique (Gaudiano et Chang, 1997). Divers avantages et lacunes existent pour chacun des modèles proposés. En général, élaborer une procédure de CO pour un robot qui est contrôlé par un RNAI s'avère une tâche ardue et requiert une puissance computationnelle élevée. C'est dans un but d'optimisation et de généralisation de cette tâche de reproduire du CO avec des RNAI et en robotique qu'un article de recherche a été produit dans le cadre de cette thèse doctorale.

À l'aide du logiciel SIMCOG, les quatre situations de base du CO ont été reproduites avec des robots virtuels et physiques contrôlés par des RNAI. Ces réalisations ont permis de dégager une invariance structurelle et fonctionnelle des RNAI. En effet, il suffit d'une cellule sensorielle, d'une cellule motrice, d'une cellule prédicteur de récompenses ou de punitions et d'un contexte d'apprentissage par renforcement pour constituer la base minimale d'un CO. Dans la Figure 4.3, les liens synaptiques sont précisés avec une synapse sans apprentissage entre la cellule prédicteur et la cellule

motrice, une synapse avec une fonction d'habituation entre la cellule sensorielle et la cellule motrice et une synapse avec une fonction STDP entre la cellule sensorielle et la cellule prédictive. À l'aide de simulations randomisées, nous avons démontré que la force des liens synaptiques peut varier légèrement sans en modifier pour autant l'issue comportementale du phénomène. Ces composantes fonctionnelles minimalistes du CO, se sont avérées suffisantes pour l'exploration de scénarios d'apprentissage variés, dans un agent complet virtuel et physique.

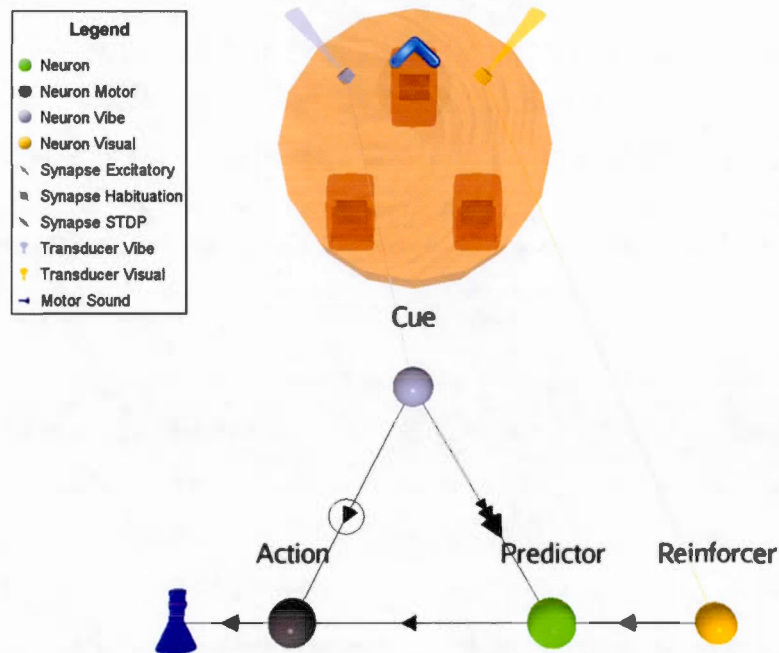


Figure 4.3 Motif cellulaire singulier capable de simuler un CO.

## ARTICLE 3

Operant conditioning: a minimal component requirement in artificial spiking neurons designed for bio-inspired robot's controller

Frontiers in neurorobotics

## Résumé

À travers un paradigme de bio-inspiration, cet article investigate le conditionnement opérant (CO) comme processus d'apprentissage pour des RNAI utilisés comme cerveaux-contrôleurs de robots. Pour les agents naturels, le CO résulte de changements comportementaux appris par les conséquences d'actions passées, influencées par divers ajustements concernant la prédiction de récompenses ou de punitions anticipées. Dans un contexte de neuro-robotique, des robots autonomes virtuels et physiques pourraient éventuellement bénéficier d'aptitudes d'apprentissage similaires, particulièrement lorsque ceux-ci évolueraient dans des environnements inconnus, imprévisibles et non supervisés. Dans cette expérimentation, nous démontrons qu'une architecture simple de RNAI suffit à soutenir le processus de CO lorsqu'appliqué à de multiples scénarios d'apprentissage. La motivation d'implémenter ce nouveau modèle de CO est basée sur les récentes littératures en sciences computationnelles et en neurobiologie, relatives aux différents modèles du phénomène. Le noyau élémentaire du CO proposé dans cet article est composé par l'association de seulement quelques neurones et synapses, mais aussi par l'inclusion originale d'une règle d'apprentissage STDP et d'habituation. Lors de différents scénarios de complexité incrémentale, nos résultats démontrent que ces composantes neuronales minimales sont suffisantes pour produire plusieurs procédures canoniques de CO. Ce module singulier vise à réaliser et implémenter plus efficacement des tâches d'apprentissages complexes par CO, lorsqu'appliqué pour des RNAI agissant comme contrôleur dans un contexte de bio-inspiration robotique.

(Voir l'annexe D pour l'article intégral)





## CHAPITRE V

### DISCUSSION

#### 5.1 Robotique bio-inspirée et apprentissage

Modéliser une IA complète et autonome en robotique bio-inspirée représente une avenue émergente dans la communauté du domaine. L'hypothèse de travail de ce paradigme repose sur la possibilité de mimer une structure ou fonction biologique par l'appariement de constituants artificiels. La simulation et prédiction adéquate de données expérimentales naturelles y est souvent le but recherché.

C'est dans cette perspective que ce sujet doctoral aborde la simulation de phénomènes d'apprentissage pour des robots virtuels et physiques avec l'habitation, le CC et le CO. En termes de réalisations, un algorithme novateur adapté aux RNAI a été développé pour chacune des trois règles d'apprentissage, conduisant à autant d'articles scientifiques. L'analyse détaillée à partir de composantes primitives neuronales artificielles ainsi qu'une méthodologie incrémentale en complexité, a permis de conclure que les robots peuvent apprendre par des mécanismes et fonctions qui sont similaires aux agents biologiques, d'un niveau cellulaire jusqu'à un niveau comportemental. Équiper des robots avec de telles fonctions cognitives pourrait permettre de meilleures capacités adaptatives dans le monde réel. La portée et les applications de cette recherche demeurent cependant à être approfondies et comparées.

Outre la réalisation de fonctions d'apprentissage biologiquement inspirées pour des robots, une visée secondaire de la thèse a été d'intégrer de façon cohérente celles-ci dans un même modèle de RNAI. À ce jour, aucun modèle computationnel n'offrait cette caractéristique de réunir et d'inclure ces trois règles d'apprentissage de base. Un troisième objectif de la thèse a été d'incorporer les RNAI dans des agents cognitifs

complets situés et incarnés, ajoutant un complément de réalisme au modèle simulé et une dimension augmentée à celle de la compréhension des phénomènes. Enfin, la validation du modèle computationnel global dans le monde réel à l'aide de robots physiques a aussi permis d'augmenter la puissance du modèle d'apprentissage, mais aussi d'articuler l'important concept de temporalité qui est commun aux trois types d'apprentissage étudiés.

L'approche de la robotique bio-inspirée opte pour le mimétisme du monde naturel dans la conception de leurs modèles. Comme autre stratégie possible en robotique traditionnelle, la méthodologie ingénierie conçoit souvent les modèles dans une optique de l'atteinte d'une efficience dans des objectifs précis. De façon générale, les chercheurs en IA ciblent d'abord une fonction, une structure ou un comportement à mimer avant d'en réaliser le mécanisme sous-jacent. À cet égard, différents niveaux conceptuels théoriques et pratiques de ces éléments peuvent être envisagés.

Par exemple, développer un patron de marche bipède pour un robot peut être implémenté par diverses solutions autant d'ingénierie que bio-inspirées. Cependant, une problématique survient lorsque les chercheurs tentent d'intégrer au modèle d'ensemble, des phénomènes connexes et hétérogènes, tels que sauter par-dessus un obstacle ou courir. Dans cette perspective, il devient pertinent pour certains chercheurs d'envisager à réunir l'ensemble des fonctionnalités dans un même paradigme d'IA et d'intégrer tous les phénomènes avec un niveau similaire de détails de la modélisation.

La robotique bio-inspirée permet justement de rassembler et d'offrir une cohérence de modélisation pour un mimétisme de structures, fonctions et comportement pour un même agent cognitif artificiel. Ainsi, la stratégie développée dans cette thèse sur le sujet de fonctions d'apprentissage simulées a été celle d'apparier des données expérimentales obtenues auprès d'animaux à systèmes nerveux simples aux comportements des robots. En ce sens, l'amalgame des fonctions dans un agent complet diffère complètement du but recherché en neuroscience, qui est de valider un concept biologique spécifique par un modèle computationnel, voire robotique. Bien qu'il puisse y avoir des avantages et bénéfices dans les deux approches, cette thèse

doctorale considère surtout l'élaboration d'une IA complète et autonome, qui simule les résultats d'un maximum de phénomènes biologiques, sans devoir toutefois démontrer qu'ils expliquent complètement les mécanismes naturels.

## 5.2 RNAI

En robotique bio-inspirée, le cerveau-contrôleur peut être envisagé par un RNAI afin de mimer les phénomènes naturels. Il s'agit d'un élément crucial dans cette thèse que de cibler le bon niveau d'abstraction des fonctions à simuler tout en permettant d'observer l'adaptation comportementale des agents artificiels tels que les données empiriques le montre pour les agents naturels. La suite logicielle SIMCOG a permis cette étude analytique détaillée, allant de la cellule neuronale jusqu'au comportement.

L'élément cellulaire constitutif d'un RNAI est le neurone. La dynamique d'un neurone de RNAI fluctue selon les variations de son potentiel de membrane qui, lorsqu'une valeur seuil est atteinte, mène à l'émission d'un PA unitaire. Les PA ne représentent qu'un des modes possibles de communication d'un agent afin de s'adapter au monde environnant. L'interface entre deux neurones est la synapse, une autre composante majeure d'un RNAI. Les synapses modulent la communication neuronale par ses propriétés de plasticité. Les synapses sont considérées par une majorité de chercheurs, comme le siège des engrammes, la trace de nos souvenirs qui selon le contexte, sont retenus (à court ou long terme) ou s'effacent.

L'impact d'une modélisation détaillée des synapses dans un RNAI demeure une avenue en plein essor dans le domaine des neurosciences computationnelles. Minimale, différents états synaptiques sont possibles (Pan et Zucker, 2009) en conjonction avec une activité-dépendante des patrons de PA pré et post-synaptiques et influencent les réseaux neuronaux, incluant l'observation de différentes formes d'apprentissage (Abbott et Regehr, 2004). Quel serait le bon niveau descriptif de la plasticité synaptique (Shouval, 2011) que l'on devrait modéliser dans un RNAI, particulièrement lorsque l'objectif est celui d'un contrôleur en robotique bio-inspirée ? Demeurer au niveau phénoménologique (mécanique interne non précisée) permet une

bonne dissociation avec les structures réelles, mais crée une distance d'interprétation des données produites alors qu'une simulation plus réaliste au niveau biophysique rapproche des données empiriques naturelles, mais au détriment d'une extrapolation du modèle vers des agents artificiels.

La question n'est pas encore tranchée et dépend beaucoup des buts recherchés et des intérêts de chacun. Dans cette thèse, les synapses ont été modélisées par des changements dynamiques détaillés, variant en fonction des patrons de PA pré (habituation) et pré-post (STDP). En ce sens, les différentes règles modulant les synapses tendent à reproduire les données au niveau phénoménologique (IA) plutôt qu'en termes biophysiques (neurosciences).

Dans chacun des articles produits, les architectures étaient volontairement simples contenant tout au plus une douzaine de neurones excitateurs et inhibiteurs avec une vingtaine de synapses, transducteurs et effecteurs. Le but recherché dans cette thèse était d'apparier les connaissances et données empiriques des trois règles d'apprentissage avec des exemples de systèmes nerveux simples et offrir une compréhension détaillée de bas niveau. Bien que les caractéristiques essentielles de chacune des fonctions d'apprentissage ont été simulées dans ce cadre précis, rien n'exclut que diverses propriétés émergentes puissent survenir avec des architectures complexes de types populationnels.

Modéliser des phénomènes cognitifs humains de bas niveau à l'aide de petits circuits neuronaux artificiels spécifiques est un premier pas, mais ne garantit en rien l'atteinte d'une modélisation de plus haut niveau. Dans ce contexte actuellement restrictif, il est difficile d'anticiper si ce modèle d'IA bio-inspiré pourra éventuellement simuler des structures telles que l'hippocampe ou le cortex cérébral et constater ultimement l'émergence de phénomènes cognitifs complexes. D'un autre côté, l'étude et la modélisation de systèmes nerveux simples, tel que celui de *C. elegans* avec ces 300 neurones et 5000 synapses, est envisageable actuellement en IA. De plus, presque tous les phénomènes d'apprentissage de base se retrouvent dans de tels microcircuits et



expliqués par plasticité synaptique (Bozorgmehr et al. 2013), justifiant la pertinence de continuer les recherches en neuro-robotique à cette échelle de grandeur.

Dans le modèle intégré de fonctions d'apprentissage de cette thèse, les RNAI représentent les contrôleurs des robots autant virtuels que physiques. Une propriété remarquable des RNAI, tout comme les neurones biologiques, est celle de communiquer dynamiquement par le biais de PA, qui eux sont modulés par plasticité synaptique. Cependant, il existe une grande variété de modes de communication neuronale dans le monde naturel autre que celui des PA, qui n'en représente qu'une des facettes possibles. Entre autres, certains neurones ont plutôt une réponse tonique et graduelle sans émissions de PA (Chalasani *et al.* 2007). Pour les types de synapses, celles dites à jonction électrique affectent différemment la communication neuronale que celle des synapses chimiques. À une autre échelle, la synchronie d'ensembles populationnels de neurones évoque aussi un autre mode de communication neuronal (c.-à-d. rythme thêta). À cet égard, les fonctions d'apprentissage développées dans cette thèse impliquent seulement un type restreint de communication entre les neurones avec un modèle conceptuel de base de RNAI incluant une modélisation simpliste de synapses chimiques et excluant des phénomènes émergents au niveau populationnel de neurones. Dépasser ces limites serait une piste intéressante à investiguer dans le cadre de la bio-inspiration robotique.

Un autre sujet à explorer liant les RNAI et l'apprentissage est lorsque le résultat d'une action (supervisée ou non) et la rétroaction des événements (contexte, action, conséquence : neutre, récompense, punition) impliquent des distances temporelles variables entre eux, ce qui réfère au problème du crédit d'assignement (Sutton, 1984) ou de récompense tardive (*distal reward problem*) (Izhikevich, 2007; Soltoggio et Steil, 2013). La durée entre les événements est généralement plus grande que la seconde, ce qui dépasse la fenêtre d'interactions des PA d'origine, surtout en ce qui concerne les fonctions d'apprentissage développées dans cette thèse. L'hypothèse d'une trace événementielle en relation avec le soutien de la dopamine et sérotonine comme neurotransmetteurs est actuellement sous investigation, autant dans le domaine des

neurosciences que dans celui des sciences computationnelles. À ce jour, aucune théorie ne peut unifier et expliquer l'ensemble des relations temporelles liant les événements externes et les perturbations internes d'un agent naturel ou artificiel.

Aussi, dans un paradigme de RNAI traditionnel, chacune des règles d'apprentissage modélisées implique que le lieu du changement par l'expérience passée se trouve directement au niveau de la synapse. Cependant, il existe d'autres alternatives non synaptiques, notamment celui du changement de la composition des canaux membranaires ou de l'excitabilité neuronale intrinsèque. En effet, dans un exemple de CC (Jones *et al.* 2003), l'augmentation du nombre de canaux sodiques dans un interneurone suffit à créer une association entre un CS et un US ! Donc, les modèles cellulaires actuels des fonctions d'apprentissage sont encore incomplets et potentiellement additifs à d'autres modèles préexistants. De plus, l'impact d'inclure dans un modèle d'apprentissage, l'effet de neuromodulateurs dans l'espace de la fente synaptique, une régulation de l'activité basale des neurones par des cellules non neuronales ainsi que l'ajout de jonctions électriques, offriraient certainement des capacités adaptatives non négligeables à développer. Enfin, le niveau de détail d'un modèle d'apprentissage s'arrête-t-il au niveau cellulaire et fonctionnel des neurones ? Y aurait-il avantage à explorer les niveaux moléculaires et génétiques ?

En résumé, à l'aide de cerveaux d'animaux simples et de préparations minimalistes *in vitro* et *in vivo*, les recherches en neuroscience ont permis d'établir les bases cellulaires de l'apprentissage et de la mémoire. Il existe des preuves suffisantes révélant les mécanismes internes par lesquels un agent naturel ajuste son comportement futur de façon observable, en fonction des événements passés. Une de ces voies d'adaptation est celle de la plasticité synaptique et la thèse complète de ce doctorat repose sur cette prémisse, avec une simulation des phénomènes au niveau cellulaire à celui d'un agent cognitif artificiel complet. Ce doctorat souligne donc, le potentiel des robots virtuels et physiques comme moyen d'incarner des modèles d'apprentissage bio-inspirée, en expliquant et en montrant la capacité de mimer des mécanismes naturels vers l'artificiel, et ce de la synapse au comportement.

### 5.3 Habituation

Dans l'article produit sur l'habituation, il a été montré que les robots peuvent s'habituer aux stimuli simples, tel que les agents naturels en sont capables, avec des changements comportementaux observables et quantifiables. C'est-à-dire qu'un ensemble de caractéristiques de l'habituation en tant que procédure, processus et mécanisme interne, a été mimé avec un degré de similitude acceptable. Cette contribution scientifique a permis d'élargir le spectre du modèle de cette fonction d'apprentissage non associative en incluant un peu plus de variables temporelles reliées aux données expérimentales du phénomène. Le modèle proposé permet dorénavant la gestion d'un plus grand nombre de motifs temporels de stimuli avec des réponses comportementales similaires aux données empiriques.

Les algorithmes traditionnels utilisés avec les RNAI en conjonction avec les robots incorporaient généralement une fonction exponentielle négative simple en regard du nombre de stimuli entrants ; le comportement observé étant alors inversement et proportionnellement réduit avec la fréquence des stimuli. L'arrêt consécutif de ces stimuli entraîne alors un recouvrement de la réponse avec une fonction inverse.

L'article souligne la prise en charge d'une habituation du comportement malgré une répétition hétérogène dans le temps des stimuli entrants. Une autre caractéristique de l'habituation qui a été modélisée est celui de l'impact inter essai de la répétition de stimuli qui offre une meilleure concordance des données empiriques avec le phénomène de potentialisation. Aussi, le ISI a été démontré déterminant dans la fonction d'habituation avec un maximum d'effet pour un ISI court et inversement.

Par exemple, l'application d'un stimulus physique de faible intensité provoque lorsque répété, une habituation du comportement. C'est-à-dire que l'action consécutive aux stimuli faiblit en fréquence et en amplitude. Cette baisse comportementale s'explique au niveau neuronal cellulaire par une dépression présynaptique entre le neurone sensoriel et le neurone moteur. La mécanique interne de cette diminution d'efficacité synaptique est complexe. La littérature pointe vers une dépendance avec

plusieurs variables temporelles reliées aux motifs de stimuli entrant, un des aspects développés dans l'article produit dans cette thématique. Par ailleurs, outre la réalisation d'un modèle computationnel plus détaillé au niveau temporel de cette fonction d'apprentissage non associative, celle-ci a été développée spécifiquement pour des RNAI. De plus, il a été possible de transférer cette fonction d'habituation vers un robot physique. La résultante est que l'adaptation comportementale des robots est similaire à celle des agents naturels, mais surtout que les changements de comportement observés, reposent aussi sur le mimétisme des processus cellulaires internes.

Toutefois, des questions subsistent sur l'habituation comme en témoigne encore l'intérêt soutenu de la communauté des neurosciences. Premièrement, l'atténuation présynaptique repose sur plusieurs cascades biochimiques, en relation avec les différents contextes expérimentaux. La contribution relative de chacune de ces voies n'a pas été modélisée en IA, probablement parce que ce mécanisme est toujours mal compris au niveau neurobiologique. Deuxièmement, la trace de l'habituation dans des circuits neuronaux complexes demeure difficilement quantifiable. Est-ce possible d'extrapoler le phénomène d'habituation dans un modèle de RNAI plus complexe ? Troisièmement, la littérature nous indique que les synapses inhibitrices sont certainement des éléments à considérer dans le mécanisme de l'habituation, car ayant une modulation différente. Quel serait l'impact d'inclure ces variantes structurelles dans un RNAI pour habituation dans un contrôleur de robot ? Enfin, les fonctions mêmes de l'habituation demeurent incertaines et théoriques. Est-ce un filtre pour éviter des stimuli répétés ? Est-ce un phénomène à la base de l'attention (Balkenius, 2000) ? Est-ce que de reproduire le phénomène de l'habituation en détail à travers un modèle d'IA contribuerait mieux à comprendre cette problématique ?

#### 5.4 Conditionnement classique

Le CC trouve ses fondements en neurosciences ainsi que des explications au niveau génétique et moléculaire. Le corrélat neurophysiologique du CC est assez bien compris



à l'aide de nombreux modèles animaux. Entre autres, le clignement réflexe de l'œil chez le lapin ou encore l'acte de se nourrir chez l'aplysie ont permis de dégager un mécanisme théorique détaillé, liant un stimulus neutre, un stimulus vital, une réponse conditionnée et une réponse inconditionnelle.

Simuler un tel phénomène à travers des agents artificiels a conduit les chercheurs du domaine à extrapoler les propriétés fonctionnelles principales de cette procédure d'apprentissage et d'en préciser le bon niveau du modèle à reproduire. En robotique bio-inspirée, cette tâche consiste à mimer le CC à l'aide de RNA ou RNAI et de fonctions d'apprentissage associatives simulées au niveau synaptique. Pour ce doctorat, le choix a été celui des RNAI et de la règle STDP comme éléments de base de l'implémentation du CC pour des robots.

Par souci de cohésion de l'ensemble du projet, l'ajout de cette deuxième règle d'apprentissage cible aussi la réplication phénoménologique du CC à un niveau cellulaire. C'est-à-dire que les composantes de bases sont des neurones et synapses artificiels comportant quelques paramètres dynamiques primaires inhérents à la catégorie des modèles computationnels de RNAI utilisés.

Outre la réalisation d'une règle non associative (habituation) et d'une règle associative (STDP) dans un même modèle synaptique de RNAI, l'implémentation commune dans des scénarios d'apprentissage en robotique constitue déjà un pas vers l'avant, compte tenu de la rareté des études dans ce domaine. Cependant, le cœur de l'article relié au thème du CC et produit dans cette thèse, propose plutôt une fonction STDP qui élargit les paramètres standards de celle-ci, en accord avec les dernières avancées des neurosciences. Ainsi, au-delà de la simple modification synaptique produite par la cooccurrence temporelle précise de PA entre deux neurones liés, ce nouvel algorithme intègre un effet de fréquence de paires ainsi qu'un effet de fréquence individuelle des PA comme règle de modification synaptique.

Cette règle STDP élargie permet une plus grande adaptation en discernant certains motifs temporels liés au monde physique. En effet, l'occurrence des PA, simulée dans un robot ou réaliste avec des animaux biologiques, n'offre pas toujours des motifs



réguliers de décharges pouvant être discrétisés et corrélés. Ces irrégularités temporelles de PA correspondants à la perception dynamique de stimuli variés et imprédictibles pointent vers un modèle d'apprentissage associatif qui peut être corrélé avec un maximum de données empiriques.

Cette règle STDP est un peu plus flexible, mais n'explique cependant pas encore toutes les données biologiques observées. Selon les théories récentes du domaine, il existerait probablement de multiples règles de modification synaptique, dépendantes non seulement de l'occurrence précise dans le temps de PA, de leur fréquence de paires et de leur fréquence individuelle, mais aussi selon que les PA soient en mode pulsée ou en volée, selon le type de neurotransmetteurs impliqués ou même selon le type de cellules neuronales. Un des facteurs responsables de cette diversité est le nombre élevé de récepteurs NMDA différents, qui selon l'expression génétique, la phase développementale et l'activité des neurones ajustent dynamiquement la composition des sous-unités, ce qui se traduit par autant de propriétés différentes de règles STDP (Sanz-Clemente *et al.* 2013). À ces facteurs observés, s'ajoute aussi qu'une bonne modélisation du phénomène STDP, devrait probablement passer d'un niveau de compréhension des modifications synaptiques cellulaires à un niveau moléculaire. Ainsi la corrélation associative trouverait ses substrats entre les différentes cascades biochimiques.

Doit-on pour autant rejeter un modèle incomplet et qui repose probablement sur une mécanique trop simple, voire erronée ? La réponse réside encore dans le but. Actuellement, reproduire du CC dans un paradigme simple est réalisable et les comportements observés autant en IA qu'en biologie sont similaires.

## 5.5 Conditionnement opérant

Le CO est une autre des formes primaires d'apprentissage associatif. Il s'agit d'un mécanisme adaptatif observé et préservé, des invertébrés primitifs tels que le ver *C. elegans* jusqu'à l'homme. Une des raisons théoriques de la conservation d'une telle

fonction réside probablement dans sa capacité à générer des solutions rapides dans un contexte à multiples variables inconnues. Ainsi, les agents biologiques tirent avantage à augmenter un comportement particulier lorsqu'il est suivi d'une récompense et inversement, à en diminuer la fréquence s'ils reçoivent une punition. De telles associations contextuelles aux actions, qui se produisent dans de riches environnements dynamiques et où le répertoire des actions possibles est grand, requièrent nécessairement une forme d'apprentissage efficace.

Bien que la dynamique précise du CO demeure encore avec plusieurs variables inconnues, lorsqu'étudiée à l'échelle cellulaire et biochimique, l'association d'une action (PA) et d'un signal de rétroaction forme tout de même la base fonctionnelle du changement de la réponse neuronale. Incidemment, les modèles analogues unicellulaires actuellement proposés orientent cette association vers la rencontre de diverses concentrations moléculaires (deuxième et même troisième messenger) (Lorenzetti *et al.* 2008). Est-ce qu'une modélisation plus fine du CO offrirait un quelconque avantage au niveau d'une implémentation en robotique ?

Une bonne raison d'inclure l'ensemble des fonctions d'apprentissage primaire dans un même modèle provient de l'étude de certains phénomènes connexes. En effet, certains chercheurs (Balkenius, 2000) ont démontré qu'un des buts de l'attention est d'aider le système neurologique dans son choix d'action. En ce sens, l'habituation, le CC et le CO participent indirectement à ce choix et pourraient même possiblement à eux seuls représenter l'essentiel du phénomène de l'attention.

Dans le même ordre d'idée, des chercheurs (McSweeney et Murphy, 2009; Lloyd *et al.* 2014) ont démontré l'implication des apprentissages non associatifs dans le processus du CO. Notamment, lors d'une même session de CO, une baisse de l'effet du renforçateur y est remarquée. Cette diminution semble être attribuée au phénomène de l'habituation. L'intégration de toutes les fonctions d'apprentissage dans la compréhension et la reproduction de phénomènes comportementaux biologiques semble donc être une voie de recherche à suivre (Salotti et Lepretre, 2008).



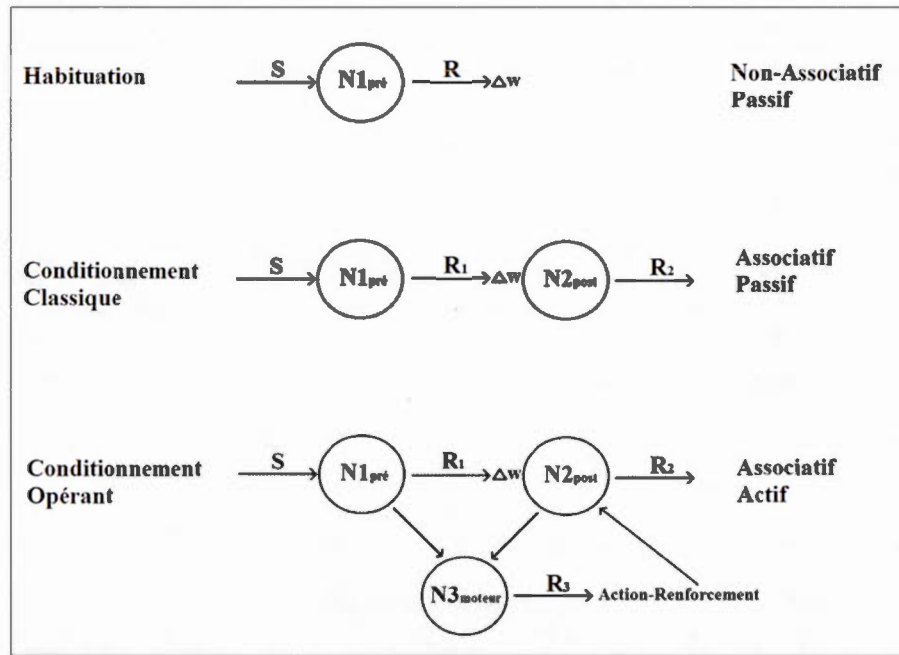
## CHAPITRE VI

### CONCLUSION ET PERSPECTIVES

L'apprentissage est une caractéristique importante de l'intelligence, permettant aux agents biologiques d'adapter leurs comportements en fonction des expériences passées. Cette capacité cognitive trouve une explication partielle au niveau de la plasticité des synapses, où les multiples transitions d'états opèrent dynamiquement par différents mécanismes neuronaux.

Pour ce doctorat, l'étude du phénomène de l'apprentissage est proposée par un modèle computationnel dans un paradigme de robotique bio-inspirée. La méthode utilisée a été de modéliser certains paramètres de l'habituation, du conditionnement classique et du conditionnement opérant, à partir de données expérimentales retrouvées chez les invertébrés. Les règles d'apprentissage proposées ont été contextualisées par de simples scénarios, sans viser des applications spécifiques. Les types d'apprentissage ont été expliqués du niveau cellulaire au comportement, par des algorithmes novateurs reposant sur une variabilité de modulations synaptiques. Ces fonctions incrémentales en complexité ont été développées et unifiées à travers un paradigme de RNAI, agissant comme contrôleurs de robots virtuels et physiques.

La réalisation de chacune de ces trois fonctions primaires d'apprentissage permet de conclure dans ce cadre précis que des robots sont capables de reproduire certains résultats au niveau phénoménologique, par des processus internes similaires aux agents biologiques. Spécifiquement, le modèle expliqué au niveau cellulaire (Figure 6.1) propose pour l'habituation et le CC, d'étendre les limites de certains paramètres dans le domaine de la temporalité. L'aspect novateur développé dans le CO a été de reproduire les scénarios de base de cette procédure d'apprentissage par un motif singulier ainsi que par les algorithmes sur l'habituation et le CC.



**Figure 6.1** Trois différents types d'apprentissage conceptualisés au niveau cellulaire.  
Légende : N = neurone, R = réponse et S = Stimulus

En résumé pour les contributions scientifiques de cette thèse :

- Intégration de trois fonctions primaires d'apprentissage dans un RNAI.
- Implémentation de RNAI comme contrôleur de robots virtuels et physiques.
- Évaluation des résultats du niveau cellulaire au niveau comportemental.
- Résultats obtenus de comportements adaptatifs temporels complexes.
  - Habituation
    - Gestion de motifs de stimuli asynchrones
  - Conditionnement classique
    - Gestions de motifs de stimuli asynchrones
    - Effet non linéaire des paires et individuel des PA
  - Conditionnement opérant
    - Motif singulier à composante minimaliste

L'impact global de ce travail de recherche en robotique bio-inspirée demeure toutefois à être plus étudié, notamment lors de situations réalistes et complexes, nécessitant un grand nombre de dimensions au niveau des senseurs, neurones et effecteurs. Les autres types d'apprentissage devront aussi s'ajouter au modèle actuel.



Enfin, la pertinence d'utiliser un modèle de RNAI et de plasticité synaptique dans la simulation de phénomènes d'apprentissage en robotique devrait également être soutenue par des études comparatives avec d'autres méthodes computationnelles.

## 6.1 Directions futures

La robotique bio-inspirée est une discipline en plein essor. Ce doctorat ne représente qu'une porte d'entrée vers ce domaine. Bien que le but ultime de l'IA soit de pouvoir un jour construire un agent cognitif artificiel possédant un maximum de caractéristiques, fonctionnalités et capacités cognitives semblables aux humains, divers chemins peuvent être empruntés pour atteindre cet objectif. Dans un premier temps, il semble opportun de se positionner dans cette communauté, en s'identifiant comme un joueur participant activement à cet objectif global et en offrant du matériel novateur dans un créneau unique. Dans cette perspective de capitalisation sur les acquis de la thèse, démarrer un laboratoire dédié à l'apprentissage en robotique bio-inspirée, en collaboration avec des universités possédant des ressources connexes à ce domaine de recherche représente un choix d'avenir logique. Réaliser un modèle intégrateur en neuro-robotique de complexité cognitive incrémentale, telle est l'ambition et la vision du projet de ce laboratoire. Accessoirement, transformer SIMCOG en logiciel libre est un objectif à venir, justifiant l'élaboration de cet outil unique tout en offrant un transfert de connaissance de cette application pour le bénéfice de la communauté. Les points suivants représentent des projets réalisables à court terme.

### 6.1.1 Évolution verticale des champs d'expertise déjà développés

Pour chacune des fonctions d'apprentissage développées dans cette thèse doctorale, diverses autres propriétés supplémentaires demeurent à être explorées dans le domaine de la robotique bio-inspirée. Par exemple, pour l'habituation, un modèle plus complet devrait inclure notamment la sensibilisation et la déshabituation comme mécanismes

complémentaires, afin de mieux refléter les observations empiriques de l'ensemble de ces formes de mémoire non associative. Dans l'apprentissage par conditionnement classique et par conditionnement opérant, de multiples concepts connexes doivent encore être formalisés pour un modèle plus complet. Également, des modèles d'apprentissage plus complexes tels que l'imitation, l'apprentissage social et même l'apprentissage de l'apprentissage (Sehgal *et al.* 2013) sont autant de phénomènes cognitifs qui devraient se juxtaposer dans un modèle computationnel complet.

Aussi, une des caractéristiques communes aux trois formes d'apprentissage qui devrait être modélisée est celle de la mémoire à long terme. En effet, la problématique est encore entière sur ce point, car les échelles temporelles utilisées actuellement dans le modèle présenté reflètent plutôt des balises entre la milliseconde et les secondes alors que dans le monde réel, les stimuli externes et les processus naturels s'étendent parfois au-delà des années. Des pistes intéressantes émanent de la communauté des neurosciences avec notamment le *tag* synaptique et celle de la convergence de concentrations d'ions calciques, mais ces hypothèses de recherche ne demeurent toutefois qu'au stade embryonnaire.

La consolidation de l'expertise déjà acquise dans le domaine de l'apprentissage est centrale dans les plans futurs et permettrait de développer divers pôles de connaissances autour de la thématique de la mémoire et de son corollaire peu encore étudié, celui de l'oubli. En général, le modèle d'apprentissage à base de RNAI trouverait aussi avantage à être systématiquement formalisé afin d'offrir des niveaux comparatifs.

### 6.1.2 Évolution horizontale des champs de la cognition en IA

Le concept d'agent cognitif artificiel complet est large de sens en IA. Celui de robot physique autonome bio-inspiré l'est tout autant et l'ajout de caractéristiques d'apprentissage n'est pas suffisant pour capturer toute l'ampleur du phénomène de l'intelligence. Maîtriser un éventail de comportements locomoteurs, exploratoires et de navigation, gérer ses ressources internes d'énergie, manipuler des symboles et

raisonner, communiquer et socialiser représentent autant d'autres objectifs à atteindre afin de constituer un modèle intégrateur d'une IA complète. Cependant, dans chacun de ces processus cognitifs, l'apprentissage joue un rôle certain, d'où la pertinence de doter minimalement un modèle d'IA avec cette caractéristique. Bien que la science traditionnelle segmente et isole les problématiques à étudier, celle d'intégrer les composantes dans un même paradigme est une tâche qui demeure à être explorée.

Par exemple, la compagnie Aldebaran robotics offre une plate-forme robotisée souple et avant-gardiste qui pourrait permettre d'inclure plusieurs de ces autres caractéristiques cognitives propres aux humanoïdes. Le prototype bipède de la taille d'un enfant nommé Nao, possède une capacité computationnelle relativement élevée (CPU / RAM) et plusieurs composantes sensorielles et motrices permettant possiblement d'atteindre des comportements plus complexes lorsqu'associés à un contrôleur adapté.

### 6.1.3 Standards de tests

Un des axes futurs qu'il serait intéressant de développer est celui de la standardisation de tests en robotique bio-inspirée, en regard de l'observation du phénomène général de l'intelligence. En effet, très peu de cadres expérimentaux, critères et balises existent en ce sens. La compétition internationale de soccer pour les robots de tout acabit est un pas bien établi dans cette direction en faisant la promotion de multiples aspects relatifs à la cognition, mais tout de même qui demeure assez spécifique. Différents autres tests inspirés de ceux en neuroscience et en psychologie sont aussi reproduits à travers divers simulateurs et laboratoires, mais aucune norme n'est encore vraiment établie dans la communauté. Ainsi, au-delà de nommer et d'inclure certaines caractéristiques de l'intelligence dans une IA, il serait opportun d'encadrer et offrir une approche systématisée de la modélisation d'une IA, à travers une hiérarchie incrémentale de tests faisant l'unanimité du domaine.

## 6.2 Autre article paru et complémentaire de la thèse

### 6.2.1 La sélection d'action

Pour les agents biologiques et artificiels, le choix d'une action plutôt qu'une autre est une forme de décision, dépassant rapidement les stades stochastiques, déterministes et réflexes comme mécanisme interne de résolution de ce problème. Ce thème est largement exploré dans la communauté du domaine sur différents niveaux conceptuels, sans toutefois encore avoir de réponse précise en regard des fondements neurobiologiques et d'un possible substrat sous-jacent à la problématique (Bryson, 2007; Humphries *et al.* 2007).

La compétition de hiérarchies structurelles neuronales semble être un de ces mécanismes de base gérant les conflits dynamiques entre les différents états internes et les stimuli externes perçus par les agents cognitifs. Est-ce possible de simuler ce mécanisme biologique à l'aide d'un modèle computationnel (Seth *et al.* 2011), d'un paradigme de RNAI (Skorheim *et al.* 2014) ou de robots ? Comment développer des architectures de RNAI et quels paramètres doivent être inclus lorsque le nombre d'actions risque d'augmenter exponentiellement tel que dans un environnement réel ? La stratégie ou l'algorithme employé pour l'implémentation systématique d'un mécanisme de sélection d'action demeure donc à être élucidé.

Différents modèles sont proposés en IA, mais encore peu sont élaborés dans une perspective spécifique de robotique bio-inspirée. Cet article tente de combler cette lacune en offrant une compréhension de bas niveau du phénomène de sélection d'action avec notamment, la démonstration qu'il est possible de mimer la fonction de sélection d'actions en incluant un motif commun et simple qui repose sur un mécanisme générateur de motifs. Ce motif singulier construit à partir d'éléments constitutifs cellulaires neuronaux primaires, permet de systématiser l'élaboration de RNAI en regard de multiples actions disponibles et même non encore apprises. Dans cette perspective, la sélection d'action reposerait sur l'influence des entrées perceptuelles et de systèmes de valeurs internes, modulés par des fonctions d'apprentissage.

En corrélat avec les études de *C. elegans* (Faumont *et al.* 2012), les décisions comportementales en contexte avec la localisation de l'habitat et les ressources de nourriture associées, des pistes sur le mécanisme de sélection d'action contenant peu d'éléments organisés en petits circuits nous sont indiquées. Notamment, des différences avec les modèles actuels sont soulignées avec l'implication de jonctions électriques (Rabinovitch *et al.* 2014) et la modulation des propriétés neuronales membranaires en sus des changements d'efficacités synaptiques. Parmi ces différents axes de solutions, un modèle de décision est suggéré dans cet article paru dans un numéro spécial de la revue *Journal of robotics: Biologically inspired robotics* (voir annexe D). L'ajout de cet article représente un complément à la thèse. Bien que non essentiel, cet article reflète bien la continuité du corpus de fonctions d'apprentissage qui a été produit et pour lequel l'intégration de celles-ci a permis d'incrémenter le niveau de complexité de la modélisation du phénomène général.



## ARTICLE 4

## Action selection and operant conditioning: a neurorobotic implementation

Journal of Robotics

## Résumé

La sélection d'action (SA) représente le processus qui est impliqué avec les agents naturels lorsque ceux-ci doivent choisir la prochaine action à accomplir. Est-ce qu'il existe un corpus élémentaire fonctionnel qui sous-tendrait ce processus cognitif ? Est-ce que l'on peut reproduire ce mécanisme avec des agents artificiels et plus spécifiquement à travers un paradigme de neuro-robotique ? Des robots autonomes non supervisés devront nécessairement recourir à un processus de prise de décision afin d'évoluer dans le monde réel et l'approche bio-inspirée est l'avenue qui est explorée dans cet article. Nous proposons de simuler un mécanisme de SA en utilisant un réseau de neurones artificiel à impulsion (RNAI) tel que retrouvé dans des microcircuits neuronaux d'invertébrés, afin d'agir comme contrôleur de robots virtuels et physiques. Ce mécanisme de SA est basé sur un simple générateur central de patrons, des neurones décideurs, des neurones sensoriels et des neurones moteurs comme composantes principales du circuit. En nouveauté, cette étude cible spécifiquement le conditionnement opérant (CO) comme contexte, ce qui est pertinent dans le processus de SA ; les choix influencent les futures rétroactions sensorielles. À l'aide d'un scénario expérimental simple, mais adaptatif, nous démontrons la complémentarité des deux phénomènes. Nous suggérons aussi que ce noyau constitutif du processus de SA représente un modèle efficient permettant de concevoir des RNAI qui pourraient inclure un nombre grandissant d'entrées sensorielles et de sorties motrices. Nos résultats démontrent que de lier un processus de SA et de CO conduit à une plus grande flexibilité comportementale lors de situations contextuelles dynamiques.

(Voir l'annexe E pour l'article intégral)

### 6.3 Articles en préparation

#### 6.3.1 Le façonnement comportemental

La poursuite logique du développement incrémental d'outils d'apprentissage en robotique bio-inspirée est de cibler une étape supérieure, et le façonnement comportemental représente une de ces thématiques dans le domaine. Cette méthodologie grandement utilisée en psychologie, permet à un expérimentateur de récompenser séquentiellement par CO, un agent biologique afin de l'amener à accomplir un objectif précis. Le façonnement comportemental capitalise sur les actions possibles du répertoire d'un agent et une modulation de leur fréquence d'apparition par une méthode de renforcement. Cette technique est appliquée dans le monde naturel lorsque les objectifs désirés ont une chance trop infime de s'accomplir aléatoirement dans une fenêtre temporelle raisonnable ou lorsque l'action visée n'est pas encore apprise par l'agent cognitif.

Les avantages en robotique bio-inspirée d'appliquer cette technique par façonnement expérimental serait hypothétiquement les mêmes que pour les agents biologiques. Ce qui éviterait une programmation lourde, redondante et spécifique pour chacun des nouveaux objectifs à accomplir, si on pouvait y palier par un algorithme qui serait généralisable.



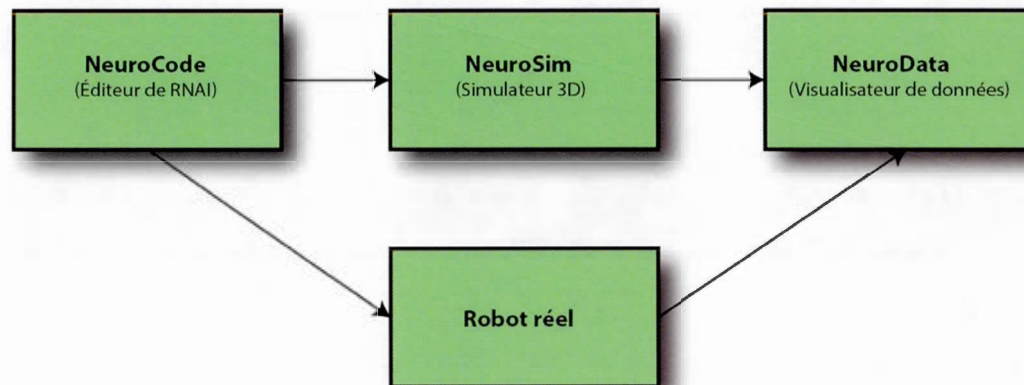
## ANNEXE A

### SIMCOG

SIMCOG est une suite logicielle conçue pour l'étude de l'IA. Ce nouvel outil dans le domaine permet de tester rapidement des hypothèses émises sur l'intelligence à travers des applications en robotique bio-inspirée. La méthodologie de SIMCOG repose principalement sur la réalisation de cerveaux contrôleurs de type RNAI, implantés pour des agents artificiels virtuels ou physiques autonomes et complets. SIMCOG est unique, car il est conçu pour simuler et analyser les processus cognitifs d'un niveau subcellulaire au niveau comportemental sur une ou plusieurs entités complètes. À l'inverse, cette suite logicielle peut également être utile pour valider des modèles en sciences cognitives ou en neurosciences, par des simulations robotiques.

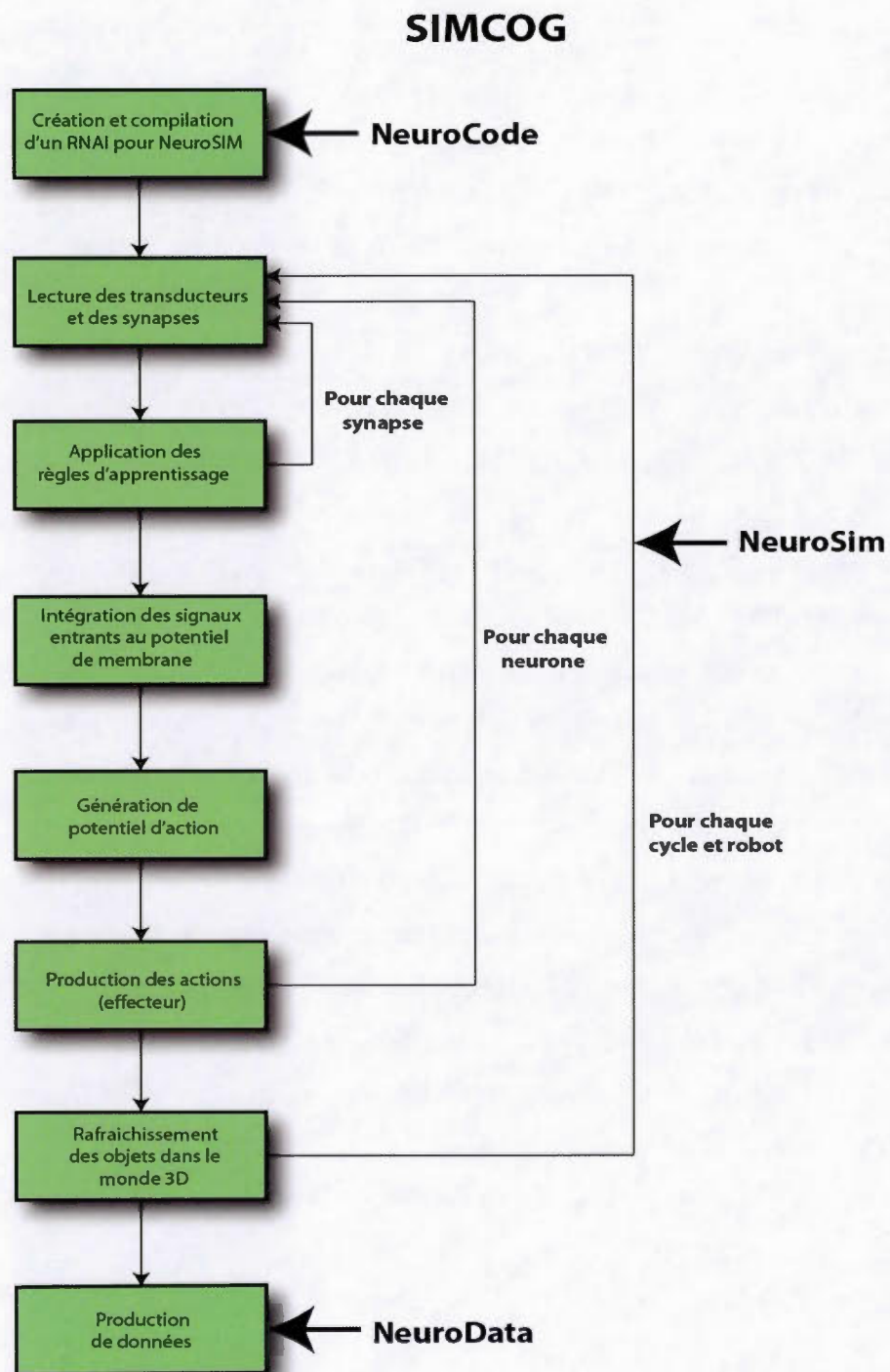
SIMCOG est actuellement composé de trois différents modules autonomes, mais interreliés : un éditeur de RNAI (NeuroCode), un éditeur et simulateur de monde 3D (NeuroSim) ainsi qu'un outil graphique d'analyse de données (NeuroData). Cette suite logicielle est réalisée par IA-Futur, une compagnie de recherche et développement en IA, en collaboration avec Objectif 8, une compagnie de développement de logiciels.

La force du logiciel réside véritablement dans la combinaison des différents modules. En seulement quelques étapes, un utilisateur inexpérimenté peut élaborer un RNAI relativement complexe et transférer quasi instantanément l'architecture compilée vers des agents cognitifs évoluant en milieu virtuel 3D ou dans un monde réel avec des robots physiques. Ce couplage fort à l'interne facilite l'automatisation de tâches répétitives telles que les ajustements de paramètres des RNAI et la validation de fonctionnalités des contrôleurs de robots. Un avantage indéniable du produit provient de sa grande convivialité des interfaces graphiques et l'absence complète de programmation à effectuer. La Figure A.1 décrit sommairement la structure de SIMCOG et la Figure A.2 résume ses fonctions par un pseudo-algorithme.



**Figure A.1** Schématisation de la structure de la suite logicielle SIMCOG. SIMCOG est composé de trois modules indépendants, mais dont les applications sont reliées. La simulation sur un robot réel est une option dans le logiciel, dont les données produites peuvent être analysées par NeuroData.





**Figure A.2** Pseudo-algorithme décrivant les principales étapes fonctionnelles du logiciel SIMCOG.

### A.1 NeuroCode

NeuroCode est un éditeur de RNAI, développé pour faciliter la réalisation d'architectures neuronales et permettant l'exploration de différentes propriétés intrinsèques des neurones biologiques. Le type de RNAI proposé dans NeuroCode est unique. Il est davantage basé sur un modèle phénoménologique plutôt que sur des équations biophysiques réalistes tel que le modèle analytique à conductance de Hodgkin-Huxley. Il s'agit d'un modèle neuronal descriptif et fonctionnel où la dynamique principale repose sur des quantités numériques relatives aux variations de courant et de potentiel de membrane. Ce modèle de RNAI bio-inspiré possède la majorité des propriétés principales des neurones ainsi qu'une cohérence d'échelle temporelle entre les différents paramètres, respectant leurs contreparties naturelles. Bien que le but recherché avec ce nouveau modèle computationnel de RNAI ne fût pas de se comparer avec les autres types de RNAI, l'élément central de l'émission de PA unique, basé sur une montée de potentiel de membrane atteignant un seuil de déclenchement, est tout de même comparable avec les modèles courants de RNAI, tel que celui proposé par Izhikevich (2003). La différence principale avec les autres modèles de RNAI réside dans son efficacité computationnelle, particulièrement lors de l'ajout d'algorithmes d'apprentissage. La composition originale du RNAI utilisé dans SIMCOG provient notamment de l'absence de calculs mathématiques complexes. Ce facteur permet d'exploiter des *central processing unit* (CPU) de faible puissance et aussi de viser des implémentations sur des plates-formes matérielles de petite taille et de faible capacité de mémoire. L'efficacité computationnelle couplée à une souplesse d'édition des différents paramètres neuronaux de NeuroCode détermine les objectifs primaires de conception du logiciel, permettant de créer des modèles de contrôleurs novateurs en neuro-robotique bio-inspirée.

Dans NeuroCode, l'éditeur est constitué d'unités de base retrouvées traditionnellement dans les RNAI et minimalement composées de neurones comme éléments centraux. Lorsque plusieurs neurones sont ajoutés dans une architecture,

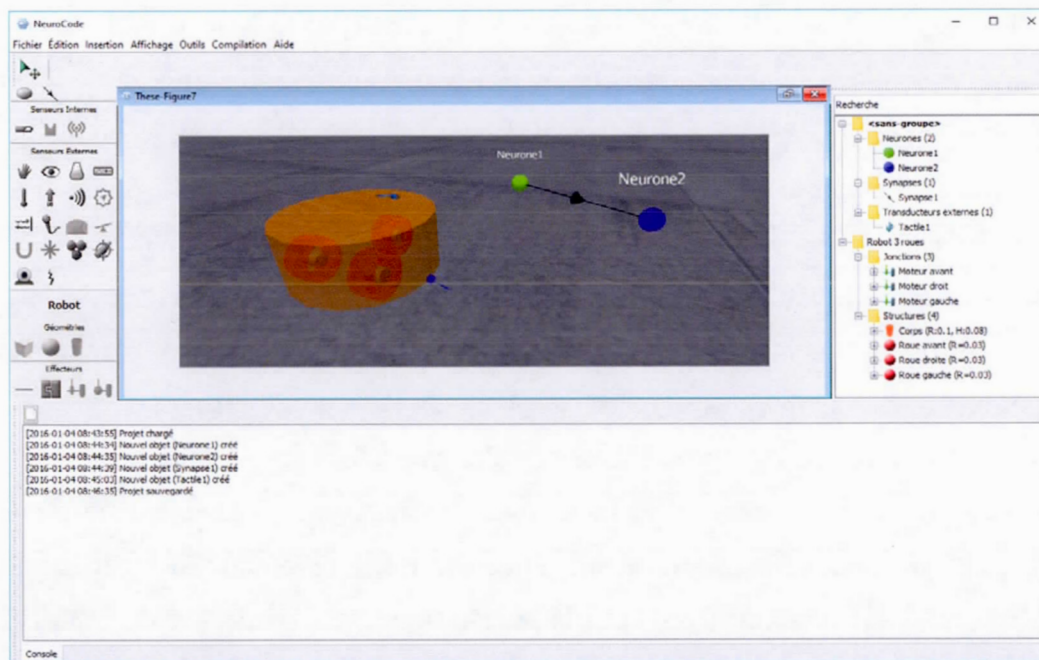
différents types de liens synaptiques permettent de les unir. Ces liens peuvent être soit uniques, autoréférentiels, bidirectionnels et/ou récurrents. La nature des synapses peut être soit inhibitrice ou excitatrice. Le niveau d'abstraction du modèle neuronal de NeuroCode cible à la fois des éléments subcellulaires tel que la gestion dynamique de fonctions pré et postsynaptiques, que l'observation comportementale de circuits complexes implémentés dans des robots complets. NeuroCode est conçu pour la gestion de RNAI à faible taille, c'est-à-dire sous une centaine de neurones et synapses dans les limites computationnelles actuelles. Dans les RNAI de NeuroCode, les neurones ne tiennent actuellement pas compte de leur aspect morphologique, donc aucun paramètre d'atténuation du courant avec la distance n'est disponible tel qu'avec les modèles de neurones à compartiments.

Un premier aspect d'innovation dans cet éditeur de RNAI réside dans la possibilité d'insérer à même l'architecture, des éléments transducteurs et effecteurs pouvant directement se connecter aux neurones. Ces deux éléments dépassent le simple cadre théorique de représenter des entrées et sorties de RNAI, en individualisant des fonctionnalités qui sont ultérieurement rattachées à de réelles composantes physiques. Ainsi, pourvus de ces éléments, les neurones deviennent qualitativement sensoriels ou moteurs. Incidemment, le type de transducteurs disponibles dans NeuroCode reflète et approxime les composantes usuelles en robotique. En outre, il est possible de définir un champ de portée de ces transducteurs. Par exemple, que ce soit des capteurs de pression, de lumière, de vibration, thermique, d'inclinaison, gyroscopique, infrarouge, ultrason, son ou caméra, les sous-composantes neuronales visent ultimement à traduire les stimuli environnementaux (virtuels ou physiques) en données numériques. Les types de moteurs varient également afin d'offrir une gamme étendue d'actions aux futures entités cognitives.

Les signaux entrants et sortants par les transducteurs et effecteurs sont aussi sujets à des modulations par des fonctions d'adaptations sensorielles et de fatigue ou baisse d'efficience, tout comme le sont les éléments biologiques analogues. Également, à l'éventail de senseurs externes disponibles, s'ajoute des senseurs internes mimant par

exemple, l'effet comportemental d'hormones artificielles ou encore, différents types de systèmes de valeurs internes qui réguleraient les RNAI.

Pour chacun des éléments de base de NeuroCode, il existe une icône dans le menu situé à la gauche de l'écran. Ainsi, un utilisateur a simplement besoin de cliquer sur les icônes et de les déplacer dans la fenêtre centrale afin de constituer une architecture spécifique de RNAI (Figure A.3). Chacun des éléments insérés est identifiable et visible en 3D. De plus, les éléments sont aisément accessibles par des outils diversifiés tels que différentes prises de vue (panoramique, orbitale, etc.), zooms et affichages sélectifs. Outre la convivialité d'utilisation, l'ensemble de ces caractéristiques esthétiques trouve écho en facilitant la compréhension d'architectures complexes.



**Figure A.3** Exemple de saisie d'écran typique de NeuroCode, démontrant les différents menus. À gauche, on retrouve les éléments constitutifs des RNAI qui peuvent être déplacés vers la fenêtre centrale 3D. Dans cet exemple, une structure circulaire de robot à trois roues, supporte un transducteur tactile fixé au bas à droite. Ce transducteur est lié à un neurone sensoriel (Neurone 2), qui lui-même reçoit un lien synaptique excitateur du Neurone 1. Chacun des éléments se retrouve dans le menu arborescent de droite, organisé en groupes fonctionnels. La séquence de création des éléments est disponible à la console dans le menu du bas. Dans cet exemple, l'utilisateur a d'abord chargé le projet à l'aide de la barre menu située dans le haut de l'écran pour ensuite sauvegarder l'image et le projet.

NeuroCode permet aussi l'insertion de structures primaires (cube, sphère, cylindre) ou complexes tel qu'un robot (c.-à-d. circulaire à deux roues), par le menu général situé au haut de l'écran ou lorsque réalisées au préalable par une importation de fichiers dans le projet en cours. Le but de cet amalgame (structure et RNAI) est de greffer les éléments y étant relatifs aux deux composantes et d'offrir un support matériel concret aux RNAI. Les avantages de cette caractéristique du logiciel sont multiples. Entre autres, la fixation des transducteurs directement sur la structure d'un robot a un réel impact sur les neurones associés. De cette façon, le codage spatio-temporel des informations devient inhérent au couple structure-RNAI, sans devoir recourir à de la programmation supplémentaire. Par exemple, si plusieurs senseurs tactiles sont distribués le long de l'axe avant-arrière d'un robot, les neurones sensoriels associés déclencheront leurs patrons de sortie en relation directe avec les stimuli perçus dans le temps et dans l'espace. Incidemment, le tandem RNAI et structure physique associée concrétise le concept bio-inspirée dans le modèle d'IA

Pour chacun des éléments de bases qui sont insérés dans l'architecture de RNAI, une arborescence visible de catégorie ordonnée par nom est construite au fur à mesure de l'élaboration du projet, dans le menu situé à la droite de l'écran. La pertinence de ce menu est de pouvoir rapidement cibler un élément en particulier, soit par son nom ou sa catégorie. C'est également par ce menu qu'il est possible d'accéder à des regroupements logiques de neurones qui auraient été préalablement constitués, tels que par exemple, des neurones moteurs consacrés à la propulsion du robot. Le regroupement d'une population de neurones offre une portée fonctionnelle et des possibilités supplémentaires de visualisation lorsque l'architecture s'avère complexe. Par exemple, un utilisateur pourrait rendre complètement invisible ou partiellement opaque certains groupes de neurones, améliorant la compréhension détaillée d'un modèle global plus complexe. Ces groupes logiques créés peuvent aussi être exportés vers une bibliothèque afin d'être importés ultérieurement.

Lorsqu'un élément est créé, un menu au bas de l'écran apparaît, donnant accès à une fenêtre de propriétés de cet item et une autre pour les propriétés visuelles générales. La

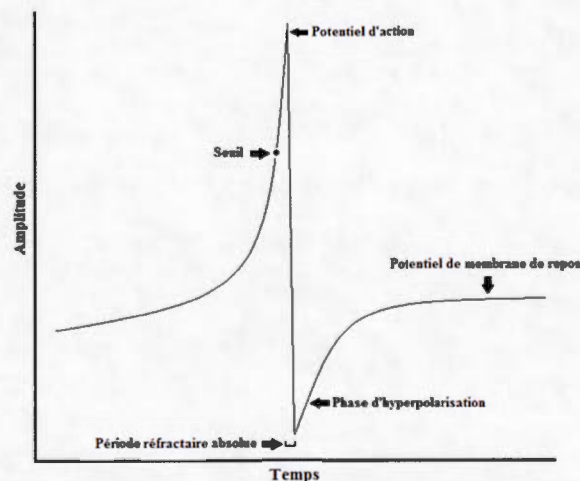


modulation des paramètres s'opère par des fenêtres à roulette avec limites numériques balisées. Une initialisation de base est faite au préalable afin de faciliter le travail d'un nouvel utilisateur. Par exemple, les coordonnées  $x$ ,  $y$  et  $z$ , les dimensions et la couleur de l'item créé sont des variables communes à tous les objets, tout comme la possibilité d'introduire ce nouvel élément dans un groupe logique.

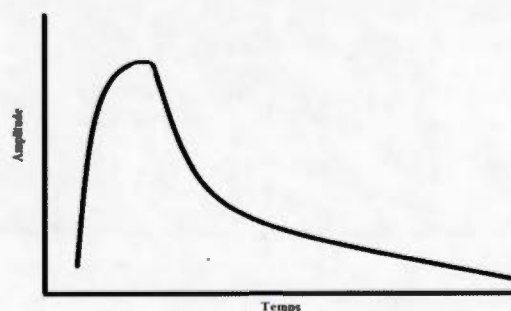
Plusieurs autres variables sont déterminantes pour qualifier un modèle de RNAI. En ce qui concerne les neurones, la variation du PM, l'atteinte d'un seuil critique de déclenchement, l'émission d'un PA, la détermination d'un potentiel de repos (PR), la durée de la période réfractaire absolue (PRA) et la fixation du retour du PM en phase d'hyperpolarisation post-émission de PA sont autant de paramètres clés qui sont inclus dans le modèle à variance de courants de NeuroCode (Figure A.4). Ainsi, un utilisateur peut individualiser chacune des valeurs numériques relativisées en pourcentage, mais aussi en changer les fonctions de base produisant les différentes courbes. Un utilisateur peut aussi ajouter des attributs stochastiques ou même des fuites de courant (fonction d'oscillation intrinsèque) afin d'engendrer des aspects dynamiques internes dans les RNAI sans devoir recourir à des entrées externes. Enfin, dans NeuroCode, chaque neurone, synapse et transducteur peut avoir des paramètres différents, pavant ainsi la voie pour la conception et l'étude de propriétés hétérogènes de RNAI.

Pour les synapses, le modèle de NeuroCode est basé sur celui des synapses chimiques naturelles. Les synapses reçoivent des signaux entrants en provenance des autres neurones et modulent ceux-ci en fonction d'une force de connexion. Un utilisateur peut initialiser cette force de connexion, mais aussi fixer un délai de transmission. Donc, lorsqu'un PA émis par un neurone A vers un neurone B est modulé par un lien synaptique S de force de connexion W, la résultante est de produire une variation locale du PM dans le neurone récepteur B. L'arrivée de ce courant entrant (PPSE ou PPSI) prend une forme graphique caractéristique de type fonction alpha (Figure A.5). Ainsi, le modèle des PPS conceptualisés dans NeuroCode est dynamique, possédant une durée et une amplitude variant selon la force de connexion et la fonction mathématique utilisée. En outre, les paramètres de cette fonction sont éditables et

individualisables pour chacune des synapses insérées dans l'architecture. Le but derrière la modélisation de ce concept est de permettre une simulation concomitante de plusieurs types de canaux ioniques et de neurotransmetteurs, afin d'engendrer des propriétés émergentes à partir de RNAI hétérogènes. C'est également au niveau synaptique que l'utilisateur peut insérer des fonctions d'apprentissage non associatives telles que l'habituation ou associatives de type STDP. Ces dernières seront discutées plus en détail dans les sections suivantes.

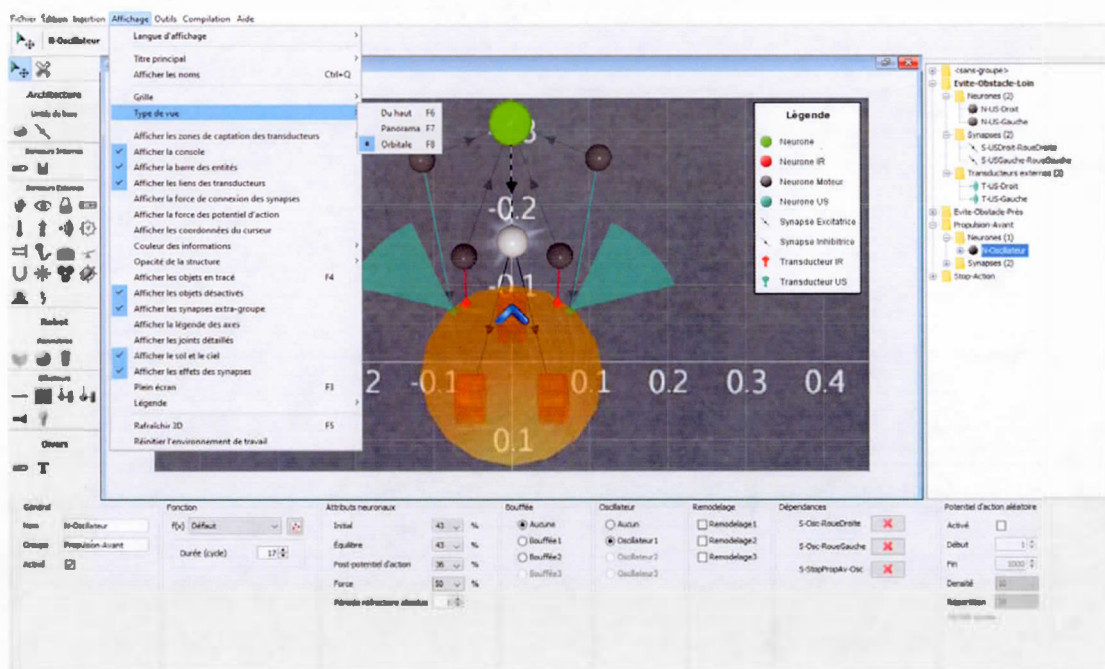


**Figure A.4** Principales composantes de la variance d'un potentiel de membrane neuronal.



**Figure A.5** Graphique d'une fonction de type alpha, représentant la variation du PM d'un neurone B, recevant un PPSE d'un neurone A et affecté d'un coefficient d'efficacité d'une synapse  $W$  reliant les neurones A et B. Les PPS sont intégrés au PM de façon non linéaire, tel que retrouvé dans les neurones naturels.

Dans NeuroCode, le menu supérieur de type barre à onglets offre à l'utilisateur une panoplie d'outils standards. Il est aussi possible d'insérer une légende ainsi qu'une grille de coordonnées pour l'organisation spatiale des architectures. L'utilisateur peut également recourir à un éditeur de texte et un éditeur de paramètres, organisés en tableaux comparatifs lorsqu'il y a beaucoup d'éléments semblables à modifier. Un outil important de ce menu est celui du générateur aléatoire d'éléments (neurone, synapse et transducteur) et de paramètres (force de connexions synaptiques). C'est aussi par ce menu qu'il est possible de compiler l'architecture de RNAI, soit vers une plate-forme physique définie ou vers un monde virtuel 3D créé à l'aide de NeuroSim. Cette option versatile est possible grâce à une imbrication de code C et Java dans le programme de SIMCOG. Enfin, toutes les actions réalisées par l'utilisateur sont visualisées par une fenêtre console apparaissant au bas de l'écran, assurant un suivi et une traçabilité du projet. Les Figures A.6 et A.7 démontrent diverses saisies d'écran représentatives de RNAI, élaborées avec NeuroCode.



**Figure A.6** Exemple d'une saisie d'écran typique d'un RNAI réalisé avec NeuroCode. Dans cet exemple, la fenêtre centrale inclut une structure de robot à 2 roues et un appui sur lequel une architecture

à plusieurs neurones, synapses et transducteurs est incorporée. Il est possible de voir le champ de lecture des transducteurs et la taille proportionnelle des neurones en rapport avec le nombre de synapses. À l'aide de la barre menu du haut, l'utilisateur a sélectionné l'onglet Affichage, puis Type de vue avec choix sur orbitale. Dans le menu de droite, chacun des éléments est individualisé et structuré en groupes logiques. Ici, le neurone oscillateur avait été sélectionné, ouvrant la fenêtre du bas en référence aux propriétés de cet objet. Dans le menu de gauche, le curseur est arrêté sur une icône de senseur externe, l'ultrason avec apparition d'une infobulle.



**Figure A.7** Exemple d'une saisie d'écran en vue panoramique réalisée avec NeuroCode. On observe un RNAI associé à une structure de robot avec différents types de capteurs et synapses. L'élément sélectionné dans le menu de droite est une synapse excitatrice avec une fonction d'apprentissage de type STDP. Dans cet exemple, l'utilisateur tente à l'aide de l'onglet Outil de la barre menu du haut, de paramétrer le générateur aléatoire pour divers éléments du RNAI.

## A.2 NeuroSim

NeuroSim est un éditeur et simulateur de mondes virtuels 3D. Ce module de la suite logicielle SIMCOG est conçu pour simuler des expérimentations portant sur l'étude du phénomène de l'intelligence. L'utilité de NeuroSim est de vérifier des hypothèses de recherche dans un monde virtuel avant l'implémentation dans des robots réels. À cet égard, il n'est pas exclu que des propriétés émergentes soient découvertes grâce à



l'élaboration de modèles simulés. Ce module se distingue de ses concurrents entre autres par son orientation spécifique vers la robotique bio-inspirée, et par ses outils uniques qui en font un véritable laboratoire expérimental dédié à l'exploration de l'IA et des grands thèmes en sciences cognitives et neurosciences tels que l'apprentissage.

Le degré d'analyse et de contrôle possible des simulations est extrême, se situant aussi bien au niveau du suivi d'un changement dynamique d'une synapse spécifique dans un RNAI, que de l'observation d'un comportement émergeant d'un groupe de robots effectuant une tâche. De plus, l'ensemble des paramètres disponibles pour moduler les simulations aspire à reproduire des paradigmes de méthodologies employées avec les expérimentations naturelles.

NeuroSim est développé avec le langage Java et utilise la librairie jMonkeyEngine<sup>2</sup> pour le rendu 3D. La physique des objets est simulée par ODE4J<sup>3</sup>, une librairie gratuite simulant la physique du monde. La séquence logique du logiciel est d'abord de créer un monde virtuel ouvert ou fermé, d'insérer des objets, puis d'exécuter la simulation dans laquelle un ou des robots interagissent avec l'environnement afin de produire des données relatives à celle-ci pour interprétation ultérieure (Figure A.8).

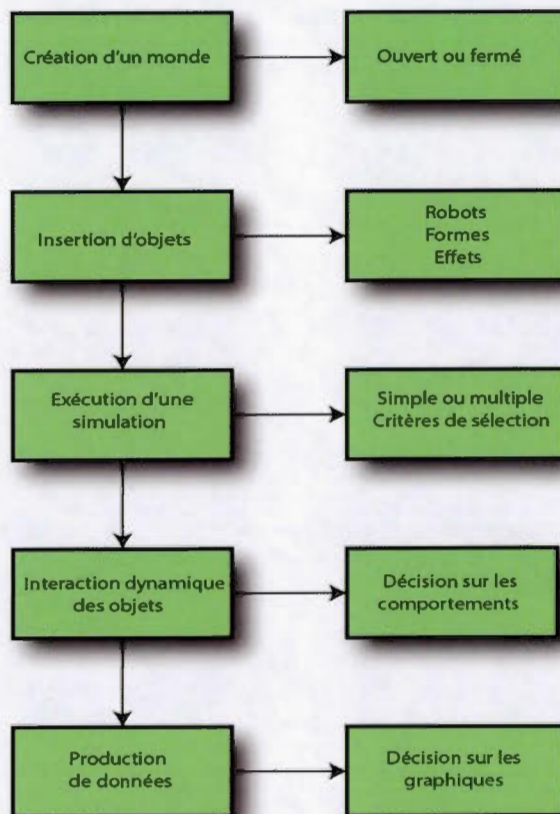
---

<sup>2</sup> <http://jmonkeyengine.org/>

<sup>3</sup> <http://www.ode4j.org/>



## Schéma logique de NeuroSim



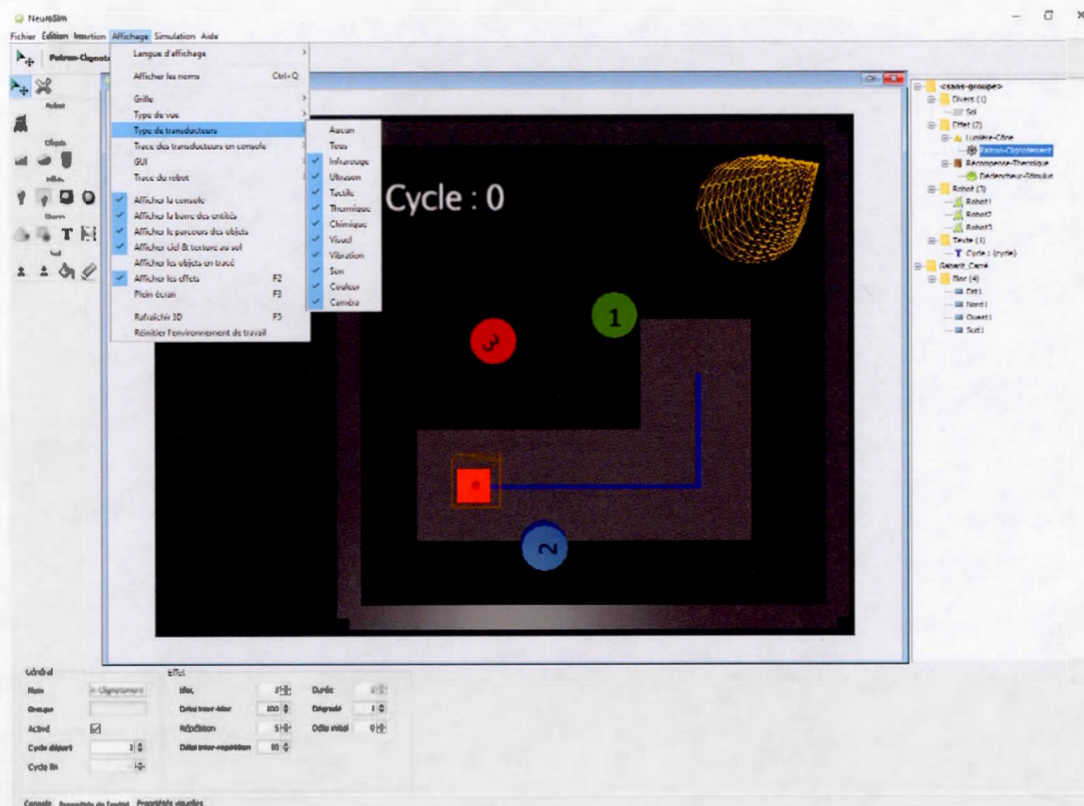
**Figure A.8** Schématisation d'une séquence logique typique utilisée par un usager dans NeuroSim.

La disposition des menus et des types d'interfaces est similaire à celle de NeuroCode. Ainsi, dans le menu de gauche, on retrouve, sous forme d'icônes, les divers objets qu'il est possible d'insérer dans la fenêtre centrale afin d'y créer un monde virtuel dynamique riche et complexe. L'élément central est généralement un robot, mais il peut également s'agir d'une structure spécifique importée par un fichier externe ou même des composantes qui seront assemblées directement dans le module de NeuroSim. Dans le logiciel, un ou des robots peuvent être insérés dans le monde virtuel et doivent être instanciés à un cerveau-contrôleur basé sur une logique de RNAI. Ainsi, un monde virtuel peut contenir simultanément plusieurs robots d'aspect structurel

différent, de fonctionnalités diverses, possédant des RNAI à caractéristiques individuelles hétérogènes.

Outre les robots, les objets disponibles de ce menu sont des cubes, sphères et cylindres. Les dimensions, positions, angles, masses, ancrages au sol, couleurs, groupes d'appartenance et diverses autres propriétés représentent autant de paramètres accessibles dans le menu du bas. Parmi ces autres propriétés intéressantes, un usager peut, par exemple, ajouter à un objet créé un ou plusieurs effets physiques et/ou un mouvement avec différentes trajectoires et vitesses.

Incidentement, la simulation d'effets physiques est un aspect novateur dans le développement de cette composante logicielle de SIMCOG. Les effets physiques sont multiples et concernent la lumière, le son, la vibration et la température. Ces effets peuvent donc être des propriétés ajoutées et liées à un objet tel qu'une sphère, un assemblage complexe d'objets dynamiques ou même un robot. Ils peuvent également être créés seuls, sans aucune référence à un objet physique. Les caractéristiques de ces effets sont soit discrétisés en stimulus simple ou continu et avoir une dispersion constante ou dégradée dans l'espace. Un patron temporel d'application des stimuli peut aussi être associé aux effets physiques, permettant de concevoir une gamme élargie d'expérimentations. Par exemple, un usager peut créer dans le coin supérieur droit du monde, un projecteur émettant une lumière tamisée et clignotante. Ce clignotement pourrait être paramétré pour une exécution en 3 blocs avec des délais interbloc de 100 cycles, où chacun des blocs comprend au total 5 répétitions avec des délais inter-répétitions de 10 cycles, pour une durée de 2 cycles par stimulus, sans effet dégradé de la lumière (Figure A.9).



**Figure A.9** Saisie d'écran typique de NeuroSim. La fenêtre centrale inclue quatre murs balisant un monde fermé, trois robots et une lumière spot dans le coin supérieur droit, associé à un patron d'effet de clignotement, tel que paramétré dans le menu au bas de l'écran. Une trace bleue au sol conduit à une zone rouge. Sur ce carré a été ajouté un déclencheur répondant à un stimulus tactile. Si un robot touche cette zone, le monde virtuel renverra au même endroit, un stimulus de chaleur. Le scénario présumé de cet exemple implique une recharge d'énergie solaire et thermique en association avec des apprentissages de localisation et de conditionnement opérant. Le curseur est arrêté sur des menus déroulants de la barre de tâche située en haut de l'écran, décrivant l'affichage de la trace de tous les transducteurs.

Un effet physique peut aussi être lié à un déclencheur non physique, celui-ci pourrait faire référence à une position précise dans le monde pour qu'un objet puisse éventuellement y interagir à un temps précis de la simulation. Un déclencheur peut aussi être un objet du monde tel qu'une tuile spécifique (c.-à-d. rouge) au sol ou une sphère que l'on touche ou applique une certaine force ou déplacement. L'action reliée au déclencheur peut se dérouler à un endroit différent de la position du déclencheur tout comme il est possible d'insérer un délai ou d'associer un patron d'effet spécifique à cette action. Également, un usager peut à tout instant, actionner dynamiquement un

bouton spécial externe pendant l'exécution d'une simulation, un déclencheur de son choix, préalablement instancié.

Comme autres objets que l'on peut insérer dans la fenêtre centrale, il y a des modifications de sol incluant, des outils pour colorer les tuiles selon le motif de choix de l'utilisateur ou même la possibilité de créer des dénivellations à l'infini pour des environnements plus réalistes. Enfin, un objet inusité de NeuroSim est celui des substances. Une substance peut être créée afin d'explorer des thématiques cognitives telles que celle des valeurs internes reliées à de la nourriture virtuelle, par exemple. Ces substances ont, entre autres, des propriétés spécifiques de patron d'évaporation et de densité, ouvrant la voie aux dimensions sensorielles olfactives et gustatives, des thèmes encore peu explorés en robotique. Cette caractéristique pourrait éventuellement trouver une résonance dans le monde physique.

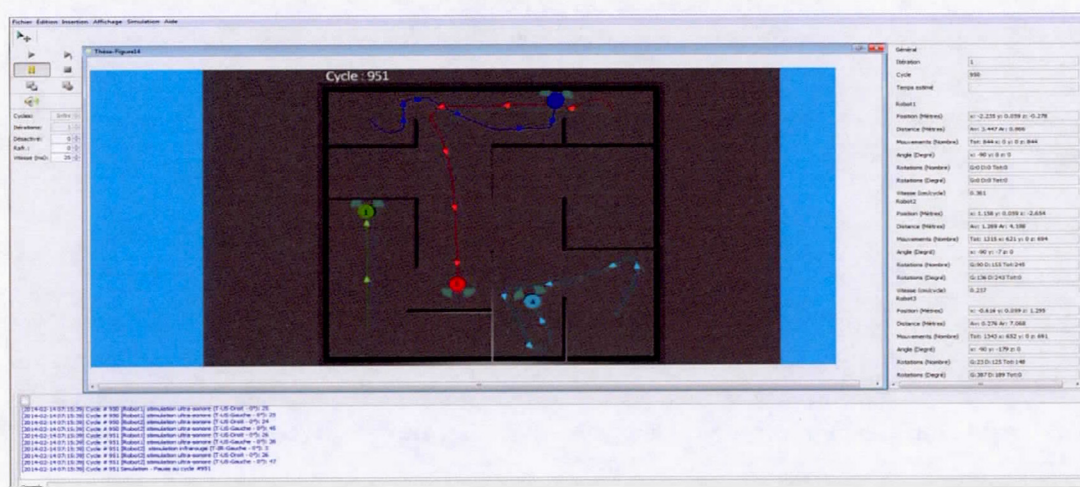
Tout comme NeuroCode, le menu de droite consiste en une arborescence des objets créés par l'utilisateur, classifiée par type et ordonnée par nom. Le menu du bas représente la console et feuille de route indicatrice des événements produits séquentiellement lors de l'édition du monde ou lors de simulations. Alternativement, le menu du bas devient la fenêtre de paramètres disponibles et éditables de l'objet créé ou pointé par l'utilisateur. La barre et menu du haut permet la gestion usuelle des fichiers et graphiques en termes d'ouverture, fermeture, importation et impression. À travers ce menu, il est possible de contrôler les types de vues, les zooms et les différents autres modes d'affichages. C'est aussi par ce menu qu'un utilisateur peut choisir d'insérer des structures préalablement assemblées tels que des labyrinthes (c.-à-d. en Y, en T, en étoile, etc.).

Enfin, un onglet de cette barre de tâches située au haut de l'écran est dédié aux caractéristiques du mode simulation. Ainsi, il existe un bouton de démarrage, de pause, d'arrêt, de mode photos ou vidéos, de téléportation dynamique d'objets, d'interaction avec des déclencheurs, de fixation du temps de simulation et de la vitesse de la simulation par cycle. Un outil intéressant de cette section réside dans la possibilité d'effectuer plusieurs simulations itératives avec ou sans modificateurs sélectifs de critères. Par exemple, cette caractéristique devient intéressante lorsqu'un utilisateur doit



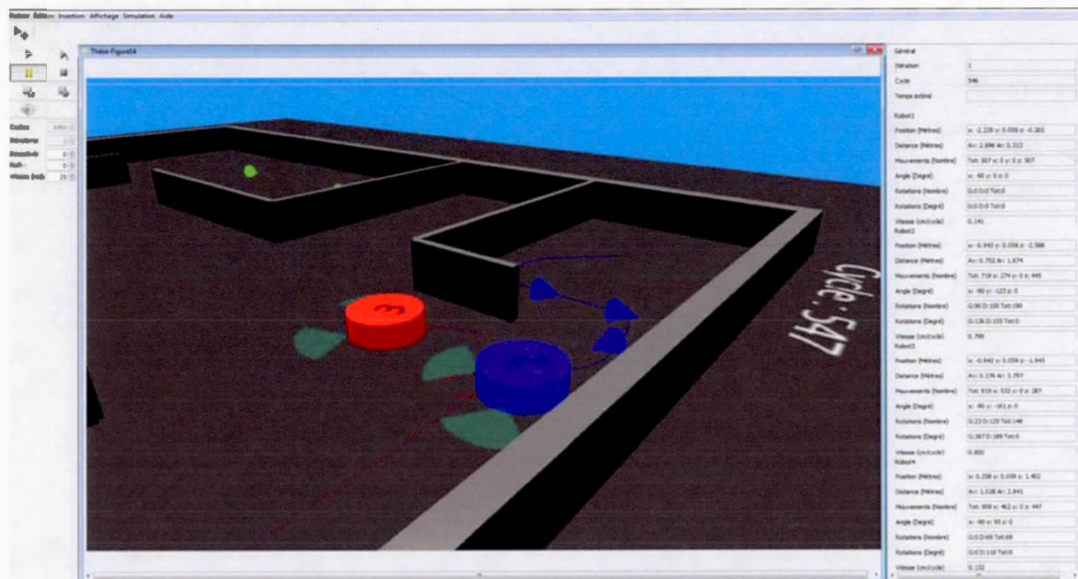
répéter des simulations avec changements aléatoires de positions de départ des objets et des robots.

Lors des simulations, les informations pertinentes aux robots sont accessibles en temps réel dans une nouvelle fenêtre située à droite. Ce qui permet à l'utilisateur de prendre des décisions d'arrêter ou non la simulation sur la base de critères observés. Par exemple, les positions x, y, z, axe de rotation et la vitesse de déplacement sont de bons indicateurs de mouvements. Une absence prolongée d'une variation sur ces paramètres en concomitance avec un robot qui ne semble pas bouger permet à l'utilisateur de rapidement cesser la simulation en cours et modifier les paramètres du RNAI et/ou du monde virtuel. Les Figures A.10 et A.11 sont des saisies d'écran représentatives de NeuroSim.



**Figure A.10** Exemple d'une saisie d'écran dans NeuroSim qui démontre un comportement exploratoire d'un labyrinthe complexe par quatre robots. L'évitement des obstacles est effectué par deux types de senseurs de proximité rapprochée (IR) et éloignée (US), visible par les champs de transducteurs des robots dans l'écran central. La trace du parcours effectué par les robots est visible. Rapidement, il est possible pour l'utilisateur de constater que le robot no.1 a de la difficulté à gérer un obstacle de front, indiquant des modifications d'architectures du RNAI à effectuer. La console au bas de l'écran indique en temps réel, les robots qui perçoivent des stimuli avec le type de senseurs et stimuli. Le menu de droite indique les données intéressantes en temps réel afin de faciliter la prise de décision de l'usager de continuer ou suspendre la simulation en cours.





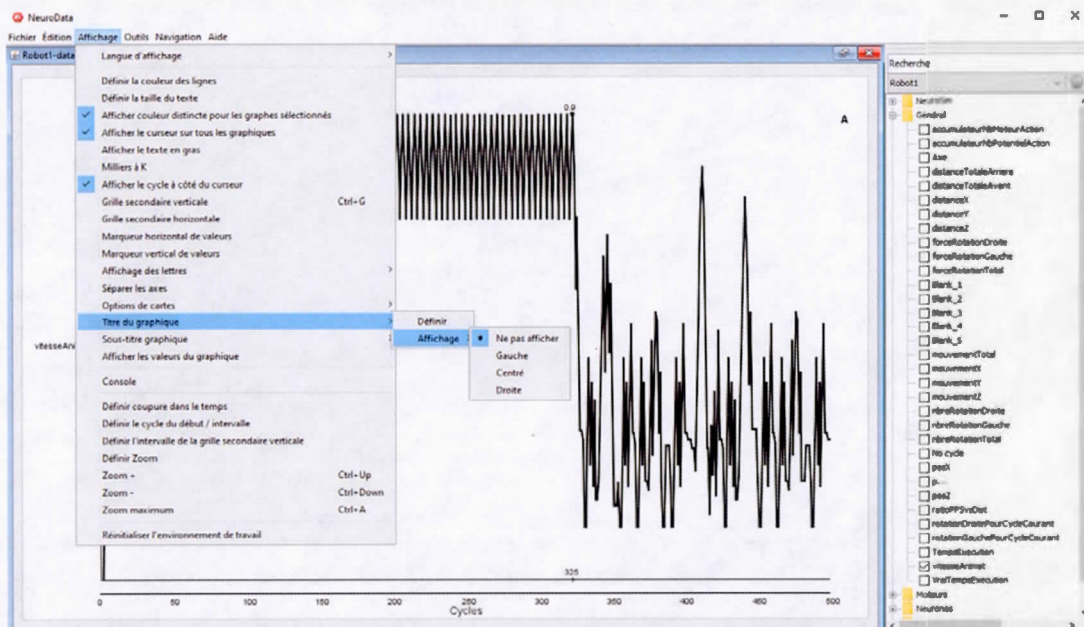
**Figure A.11** Exemple d'une saisie d'écran du même projet que la Figure A.10, mais avec une vue orbitale rapprochée, permettant de mieux apprécier le champ des transducteurs, la direction des flèches du chemin parcouru par les robots et les possibles collisions entre les robots et les murs du labyrinthe.

### A.3 NeuroData

La composante logicielle NeuroData a été conçue dans le but spécifique de produire des représentations graphiques à partir des données issues de simulations exécutées dans NeuroSim ou pour les robots réels. Tout comme les autres composantes de SIMCOG, l'aspect convivial de l'interface de NeuroData est un atout important. Le côté pragmatique et appliqué est aussi souligné en rapport avec des sorties de fichiers destinés à un usage d'insertion dans des articles pour des revues scientifiques.

De façon générale, un usager ouvre le fichier correspondant au nom d'un robot qui est associé à une simulation, à l'aide d'un onglet dans le menu supérieur. Le contenu de ce fichier apparaît ensuite dans le menu arborescent situé à la droite de l'écran, où figurent toutes les variables disponibles à visualiser, classifié par groupements logiques, types d'objets et ordonnées alphabétiquement (Figure A.12). Fait à noter, ces variables incluent aussi bien des paramètres du RNAI du robot que ceux de la gestion interne et externe des objets de NeuroSim pour les simulations effectuées. Par exemple,

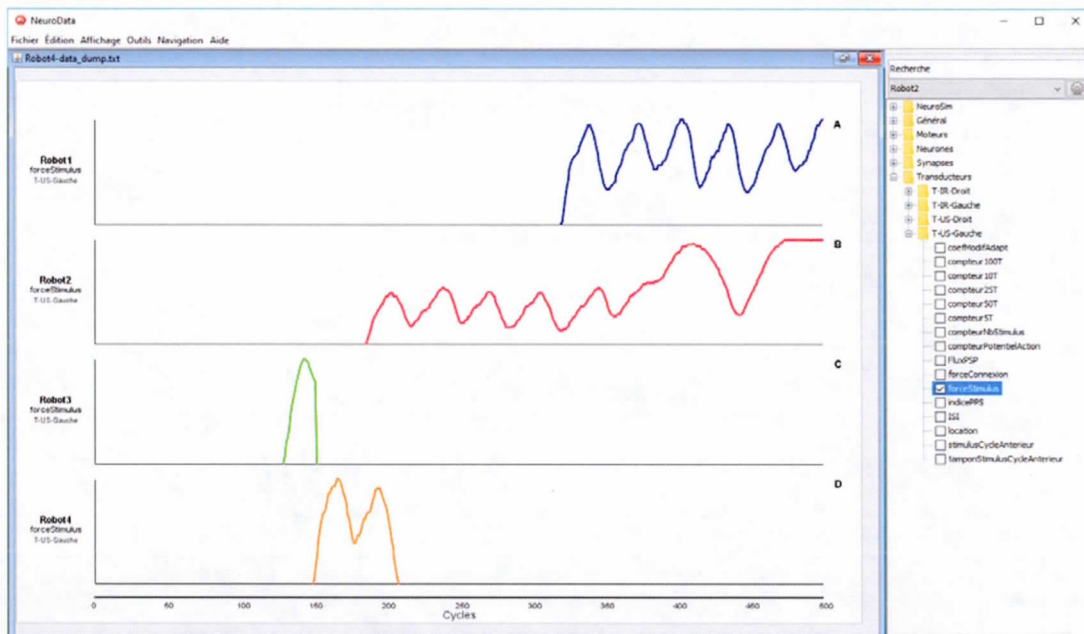
il est possible de visualiser dans un même graphique le potentiel de membrane d'un neurone J, la vitesse de simulation par cycle exécuté dans NeuroSim, le pas de déplacement sur l'axe des X d'un robot R et l'état de l'intensité d'une lumière L dans le monde virtuel. En cliquant sur un objet dans le menu de droite, un graphique apparaît dans la fenêtre centrale, correspondant aux données numériques y étant reliées selon l'ordre chronologique des cycles de la simulation.



**Figure A.12** Exemple d'une saisie d'écran typique de NeuroData. L'utilisateur a d'abord ouvert le fichier Robot1 d'une simulation et cliqué dans le menu arborescent de droite sur une variable disponible. Dans cet exemple et en référence au monde virtuel des Figures A.10 et A.11, la vitesse de déplacement a été choisie avec la visualisation de cette variable sur une fenêtre de 500 cycles. C'est peu après le cycle 325 (voir marqueur vertical sur l'abscisse où la valeur correspondante de la vitesse est de 0.9 cm/sec) que le robot1 a dû rencontrer le mur et décélérer. À l'aide du menu du haut, l'utilisateur a choisi l'onglet Affichage, puis Titre graphique afin de centrer le titre choisi.

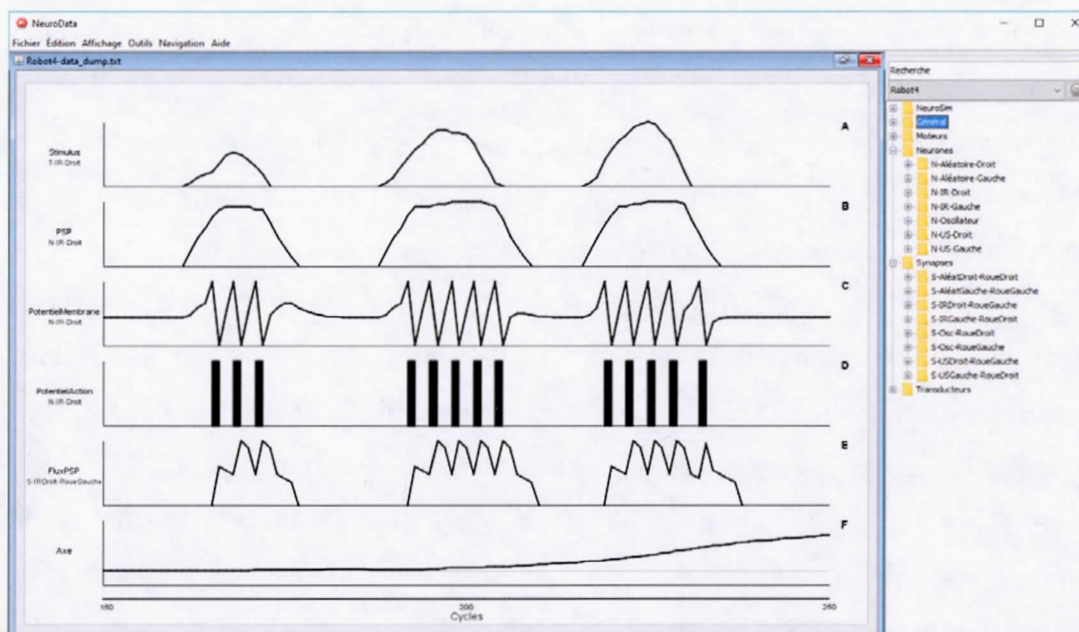
Les outils disponibles sont multiples dans NeuroData tels que des zooms, des fonctions d'éditations de texte, de paramètres concernant le trait des courbes, grilles, axes, marqueurs d'axes, échelles et titre. Une page graphique typique peut contenir plus d'une courbe à la fois afin de permettre des comparaisons simultanées et des suivis spatio-temporels sur plusieurs variables. Il est également possible de comparer des

variables en provenance de robots différents dans une même simulation ou même pour des simulations différentes. Enfin, une fonctionnalité intéressante de NeuroData est celle de pouvoir visualiser en temps réel les données produites soit par les simulations virtuelles de NeuroSim ou même avec les robots réels. Les Figures A.13 et A.14 sont représentatives de saisies d'écran de cette composante de la suite logicielle SIMCOG.



**Figure A.13** Exemple d'une saisie d'écran en rapport avec la simulation des Figures A.10 à A.12, qui démontre une comparaison d'une même variable dans un même graphique pour les quatre robots de la même simulation. Dans cet exemple, la variable choisie est celle du stimulus capté à l'aide des transducteurs à ultrason de gauche.





**Figure A.14** Dans la même simulation que les Figures A.10 à A.13, saisie d'écran démontrant une séquence logique temporelle pour le robot4. D'abord, il y a captation du stimulus par le transducteur (A) et conversion de ce stimulus en courant entrant, causant une dépolarisation locale et isolée en absolu (B). Ce potentiel récepteur fait varier le potentiel de membrane du neurone (C) qui selon le seuil déclenche une série de potentiels d'action (D). Ces potentiels d'actions traversent une synapse et selon la force affectent le courant (E). Cette synapse est associée à un effecteur, convertissant le courant en potentiel moteur. L'activation du moteur-roue occasionne un changement rotatif du robot qui est visualisé par le tracé de l'axe (F).

#### A.4 NeuroSync

Élaborer une architecture de RNAI complexe en vue de résoudre une tâche cognitive ou mimer un comportement biologique à travers une expérimentation sur des robots virtuels ou physiques demande souvent une collaboration entre plusieurs individus. L'atteinte d'objectifs supérieurs en IA et en robotique cognitive exigera à coup sûr le travail par de multiples équipes. Travailler derrière un écran à plusieurs intervenants dans un même lieu physique est parfois laborieux, voire souvent non réaliste. Ainsi, le morcellement et la répartition du travail sont inévitables avec un besoin d'orchestration sans toutefois perdre de vue l'ensemble des particularités de chacun. À l'ère numérique,

de l'internet et du travail à distance, SIMCOG a su s'adapter à cet incontournable paradigme méthodologique.

NeuroSync est un outil supplémentaire actuellement disponible dans SIMCOG, permettant à l'aide d'avatars, d'offrir la possibilité à plusieurs utilisateurs de construire ensemble et simultanément des RNAI, peu importe leur lieu de travail.

Bien que l'idée de NeuroSync soit déjà assez novatrice et pratique, l'arrivée des réalités virtuelles augmentées précipite aussi l'exploration plus avant de cette nouvelle technologie. Notamment, par le biais d'une interface avec des lunettes à écran 3D, tel que la compagnie Oculus en produit. Dans cette perspective, non seulement NeuroSync permettrait à différents utilisateurs de créer ensemble des RNAI et des mondes virtuels complexes, mais cette création pourrait se faire de manière libre, sans clavier par l'inclusion complète des avatars dans le logiciel. Ainsi, la manipulation de palettes d'outils et des objets s'effectuerait directement dans un monde virtuel. Cette avenue est actuellement sous la loupe et en discussion.

#### A.5 NeuroGene

Les RNAI tels que modélisés et utilisés dans la thèse semblent être de niveau conceptuel adéquat pour résoudre des problématiques de base en apprentissage pour la robotique bio-inspirée. Cependant, force est de constater qu'il y aura toujours un goulot d'étranglement avec la complexification de l'architecture neuronale. Par exemple, lors d'ajouts de propriétés neuronales ou de comportements à mimer, le nombre de variables pourrait rapidement atteindre l'explosion combinatoire.

Ainsi, l'étape manuelle d'ajustement de multiples paramètres est un réel fardeau dans le développement d'un contrôleur de RNAI complexe, tel que celui permettant de résoudre un ensemble de tâches diversifiées et inhomogènes. Traditionnellement, les chercheurs utilisent diverses méthodes computationnelles afin de pallier à cet obstacle. Les AG représentent une de celles-ci en exploitant des critères de sélection basés sur les fonctions des processus génétiques et du moteur évolutif. Cette méthode pourrait



être une alternative bio-inspirée pour le paramétrage des RNAI (Cachón et Vázques, 2015) tel que celui du codage initial des poids synaptiques.

D'ailleurs, il existe des solutions efficaces qui incorporent des algorithmes évolutionnaires comme méthodes de variation de paramètres dans les RNAI (Carlson *et al.* 2014, Fontana *et al.* 2014). Cependant, il serait opportun d'évaluer SIMCOG en ce sens, avec l'ajout d'un module supplémentaire dédié aux AG. À ce jour, NeuroGene est à l'étape de développement avec toujours comme objectif l'innovation, la convivialité d'utilisation et son intégration fluide avec les autres modules.

Deux autres raisons de joindre des AG aux RNAI en robotique bio-inspirée sont la réalité biologique des changements synaptiques qui reposent sur un mécanisme dynamique de communication avec le matériel génétique (LTD/LTP) ainsi que tout le champ d'expertise de l'épigénétique. Enfin, cet outil pourrait permettre d'explorer un autre champ en IA et en robotique bio-inspirée, celui de l'aspect du développement, à tout le moins le développement structurel des RNAI basés sur des AG (NeuroDev). Cette voie de recherche devient particulièrement intéressante lorsqu'elle inclut en corrélation, les études neurobiologiques actuelles sur le développement des processus d'apprentissages du CC et du CO (Valente *et al.* 2012).

Utiliser des fonctionnalités génétiques et développementales en robotique bio-inspirée permettrait aussi d'explorer un champ émergent en AI, celui de la robotique évolutionnaire (Doncieux *et al.* 2015; Nolfi et Floreano, 2000; Bongard, 2013). Évoluer des robots grâce aux principes de variation et de sélection (structurelle et fonctionnelle) représente à mon avis, la seule avenue valable pouvant conduire à s'approcher un jour du mimétisme de la cognition humaine.

Enfin, que ce soit un contrôleur de type RNAI, une fonction d'apprentissage, une structure robotique avec déclinaison de différents senseurs et effecteurs, le parallèle dans le monde naturel est que l'ensemble des informations conduisant à l'élaboration de l'agent complet se situe sur un brin d'ADN et un contexte environnemental précis. Développer une IA complète à partir de l'inspiration biologique devrait donc logiquement conduire à simuler un jour ce *blueprint* pour réaliser ce modèle d'IA.

## ARTICLE 5

AI-SIMCOG: a simulator for spiking neurons and multiple animats' behaviours

Neural computations &amp; applications

## Résumé

Construire une architecture neurale bio-inspirée pour un contrôleur de robot afin de tester des hypothèses en cognition et sur l'intelligence n'est pas trivial, surtout si ce contrôleur est un réseau de neurones artificiels à impulsions (RNAI). En conséquence, les simulateurs intégrant des RNAI à des environnements de mondes artificiels ou à des robots physiques sont peu nombreux. Dans cet article, nous présentons un simulateur d'intelligence artificielle permettant d'exploiter la force computationnelle des RNAI lorsqu'associée à de petits systèmes cognitifs artificiels. Cet article souligne la convivialité d'utilisation d'interfaces graphiques, les simulations en temps réel et la modélisation de synapses dynamiques adaptatives pour les RNAI. L'inclusion de transducteurs et de composantes hormonales, d'oscillateurs intrinsèques ainsi que différentes règles d'apprentissage incluses dans un algorithme discret de RNAI représentent autant d'aspects distinctifs de ce logiciel. Une caractéristique additionnelle majeure est celle d'un lien efficient entre les architectures de RNAI et le simulateur de monde 2D, permettant l'implémentation d'un ou plusieurs animats en temps réel. Incidemment, avec ce logiciel, il est possible de transférer directement l'architecture de ces contrôleurs vers des robots physiques. Enfin, ce modèle neural computationnel permet d'investiguer plusieurs aspects temporels reliés aux divers problèmes en cognition.

(Voir l'annexe F pour l'article intégral)

## ARTICLE 6

NeuroSim: a virtual 3D world to investigate the intelligence phenomenon within the perspective of bio-inspired robotic agents

Chapitre dans le livre : Virtual world

**Résumé**

Cet article décrit NeuroSim, un nouveau simulateur dans la communauté de l'intelligence artificielle. Composante intégrée de la suite logicielle SIMCOG, NeuroSim a été développé pour l'investigation du phénomène de l'intelligence générale en utilisant une approche de robotique bio-inspirée. Par une importation d'architectures composées de réseaux de neurones artificiels à impulsions agissants comme cerveaux-contrôleurs de robots, NeuroSim permet d'étudier en détail les comportements reproduits dans le monde virtuel 3D. Dans cette perspective, plusieurs types de vues offrent des observations en temps réel et permet une production massive et détaillée des données de sortie, aidant les utilisateurs à mieux comprendre les résultats de leurs simulations. Au-delà de la création de robots virtuels et d'objets primaires ou de canevas de labyrinthes qui peuvent être ajoutés dans l'environnement, l'originalité de ce simulateur consiste avant tout à intégrer les stimuli en provenance de multiples modalités sensorielles ainsi que l'application de patrons temporels complexes d'effets physiques simulés, pour des contextes variés d'étude de l'apprentissage.

(Voir l'annexe G pour l'article intégral)



## ANNEXE B

### ARTICLE 1 : HABITUATION

#### BIOINSPIRATION & BIOMIMETICS



# Habituation: a non-associative learning rule design for spiking neurons and an autonomous mobile robots implementation

André Cyr and Mounir Boukadoum

Computer Science Department, Université du Québec à Montréal, Montreal, Quebec, Canada

E-mail: [andre.cyr1@videotron.ca](mailto:andre.cyr1@videotron.ca) and [boukadoum.mounir@uqam.ca](mailto:boukadoum.mounir@uqam.ca)

Received 5 August 2012

Accepted for publication 9 January 2013

Published 5 February 2013

Online at [stacks.iop.org/BB/8/016007](http://stacks.iop.org/BB/8/016007)

## Abstract

This paper presents a novel bio-inspired habituation function for robots under control by an artificial spiking neural network. This non-associative learning rule is modelled at the synaptic level and validated through robotic behaviours in reaction to different stimuli patterns in a dynamical virtual 3D world. Habituation is minimally represented to show an attenuated response after exposure to and perception of persistent external stimuli. Based on current neurosciences research, the originality of this rule includes modulated response to variable frequencies of the captured stimuli. Filtering out repetitive data from the natural habituation mechanism has been demonstrated to be a key factor in the attention phenomenon, and inserting such a rule operating at multiple temporal dimensions of stimuli increases a robot's adaptive behaviours by ignoring broader contextual irrelevant information.

(Some figures may appear in colour only in the online journal)

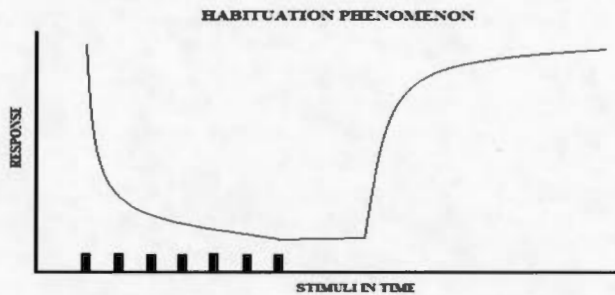
## 1. Introduction

A recognized hallmark of natural intelligence is the capacity of an organism to adapt to its environment. Even for simple organisms, adaptation means coping with internal and external constraints in a complex and dynamic real-time environment. Adaptations can be small or big, transient or permanent, and can be explained on many different conceptual levels. Several adaptation mechanisms are known, but learning is probably one of the most helpful during the lifetime of natural organisms. One of the primary universal low-level mechanisms of learning is habituation, a non-associative process of adaptation [1].

Traditionally, habituation has been understood from two viewpoints: psychological and neurobiological. As a result, it has been studied through animal behaviour, animal learning and nervous system perspectives. A convenient definition for both viewpoints represents habituation as a temporary gradual down-modulation of the intensity and frequency of

the behavioural response to repetitive stimuli or events. The main characteristics of the habituation process were detailed in the late sixties [2, 3] and recently revisited by Rankin *et al* [4].

Habituation has been explored in several ways in natural organisms. Variation of the presynaptic element is a major factor to explain the non-associative learning function validated through the synaptic plasticity phenomenon, hence investigations involving experiments on synapses and single neuron [5], neural circuits [6] and even well-defined brain structures [7]. Moreover, habituation has been demonstrated in almost every sensorial modality and in many living creatures, from lower organisms such as worms [8] to higher ones such as rats [9] or humans [10]. In fact, habituation is a fundamental feature of the memory processes and is linked to other learning mechanisms [11–16]. Understanding the foundation of habituation in different species and across sensorial modalities can advance our knowledge about learning and



**Figure 1.** The temporal sequence of the habituation mechanism: if repeated stimuli are applied, then an exponential decay of the normal response is observed. If the stimulus ceases, after a certain time window a recovery phase is observed, where the exponential positive slope is not identical to the previous negative decay phase.

memory, two major mechanisms that trigger adaptation in natural intelligence agents.

The properties of habituation are numerous. It is a non-associative process where the stimuli are unconditional, automatically triggering a constant response, unrelated to a secondary post-response event. Also, the stimuli should not be biologically relevant or vital to the organism; otherwise, it would lead to a fixed-pattern response for survival purposes. Instead, the stimuli should be novel and innocuous. In this sense, habituation may be a key point in arousal and attention, serving higher general cognitive mechanisms. It may act as a sensory gate and as a filter, by clearing non-essential or insignificant stimuli and indirectly highlighting novel environmental features that may be meaningful to learn for the organism. Habituation could help discriminate, segment, and in certain ways, categorize the stimuli. The habituation process is also stimulus specific [17, 18], but generalization is possible in specific contexts [19].

In addition, habituation is also locus specific [6, 20] and depends on the stimulus modality, the modality's field width, neuron properties and the complexity of the circuit involved. Habituation can also be obtained from different concomitant input modalities and multiple stimuli [21]. Furthermore, it depends on the exact developmental stage of natural organisms [22]. Finally, habituation manifests itself differently in animal phylogeny, increasing its complexity and capacity from invertebrate to vertebrate animals.

Nevertheless, habituation refers minimally to the observation of two main phenomena that occur in sequence (figure 1): an exponential decay of the response to repeated stimuli followed by a recovery phase. After the stimulus disappears or changes, a recovery phase starts as a partial return to the original baseline response, but it incurs a certain delay during which a positive exponential function is observed, with a time constant that is not simply the opposite of the negative exponential observed during habituation. If the recovery phase is short, it refers to the short-term habituation mechanism, a transient minute memory thought to be a modulation of the presynaptic element; if habituation goes on for several hours, one or many intermediary-term mechanisms occurs with the implication of translation proteins (not RNA synthesis). If the habituated response is still present after 24 h, it is classified

as long-term or long-lasting process, a relatively permanent memory that relies on the production of new materials from DNA transcription [23–25]. These different routing patterns reveal a habituation hierarchy that is based on multiple temporal features of the stimuli and relies also on the timing of the recovery phase [26, 27].

Consequently, the stimulus temporal dimensions (stimulus length, interval stimulus, inter-stimulus interval: ISI, interval training and interval variability) influence the habituation response through several pathways. These relationships characterize the habituation mechanism from a behavioural to genetic perspective [28–31]. Rankin and Broster [32] showed that proper calibration of the experiment parameters with precise ISI could lead to an accurate prediction of the recovery phase. To summarize [33]: the shorter the ISI is, the more significant the habituation response will be and the faster the recovery is from the initial baseline response. As a result, high frequency stimuli patterns lead to more habituation than lower frequency inputs. Also, the longer the ISI is, the longer the memory trace will be. Short ISIs in habituation and recovery processes are not only faster, they are also stronger and more complete than with longer ISIs. Consequently, the recovery phase of habituation depends on the exact time window of the frequencies and magnitudes of the stimuli. From a computational viewpoint, the ISI represents a useful indicator, giving the opportunity to reproduce many functional aspects of the natural habituation mechanism, including the distinction between short-term, intermediary and long-term habituation. Another observation from natural experiments is that habituation can be accumulated between trials. This persistence can be regarded as a memory trace or a potentiation of habituation.

Many authors now view habituation as multiple processes triggered from the integration of stimuli amplitude, frequency and timing, number, types, location and context [34]. Moreover, individual neurons exhibit their own kinetic signatures for the habituation process and recovery phase [33], a fact that dramatically increases the adaptation of small neural networks. Therefore, habituation could be at least explained computationally as a stimulus capture device with several time-scale parameters [35, 36] or a kind of signal-filter process [37]. In this respect, different ISIs and stimuli pattern protocols lead to dividing the habituation mechanism into a hierarchy within multiple memory phases [38] and different time courses. A general observation emerging from this brief survey of researches on habituation is that this simple learning rule is sustained by complex mechanisms covering many conceptual levels, from synaptic changes to the manifestation of global behavioural adaptations.

Different models in artificial intelligence (AI), with varying levels of abstraction, have been used in attempts to understand the habituation phenomenon. Biologically inspired approaches [39–41] are common to explore and study natural learning mechanisms. For instance, many abstract neural models have been employed in the artificial neural network (ANN) paradigm used in the AI community. This paper favours the artificial spiking neural network (ASNN) [42–45], a more biologically realistic subset of ANN that puts emphasis on



information timing, a process whose resolution is at the level of the neural spike instead of traditional rate coding. When used as controllers of AI agents possessing sets of virtual or physical sensors and actuators, ASNN models allow a bigger abstraction of their biological counterparts with the integration of input stimuli and the output signals to produce behaviours. Depending on the architectural complexity and details, these models permit different types of temporal computation.

Learning functions in conjunction with ASNNs have recently been a focus of attention [46]. Using an ASNN as a controller for robots or animats is also generating growing interest among AI scientists [47]. Less common are efforts to combine properties of natural learning process within an ASNN, embedded in physical or virtual AI agents. Previous researches on computational models of habituation have demonstrated specific biological learning features [48–54]. Some research proposes a model with traditional but dynamic ANNs [55–57], others use an ASNN, and still others embed habituation within an animat [58] or physical robot platform [59–61]. As examples, Chang [60] demonstrates the pertinence of applying the habituation rule to optimize robotic hallway navigation, but using a traditional ANN and fixed temporal stimuli (obstacles in relation to short term IR readings). Marsland [55] used a computational neural model of a hierarchical habituation rule for real mobile robot, but failed to show all the temporal features. All these works were applied in relatively simple conditions, and none of them has reached a level of complexity that reflects the current comprehension of the natural habituation model, with a view of exploiting the multiples temporal features of the learning function for autonomous mobile robots.

Also, many scientists have sought to enrich the theoretical modelling aspects of habituation by going beyond mere neuronal properties to encompass the brain, body and environment triad. In this perspective, habituation is defined through the observation of an embodied and situated agent behaving in its world [62–65]. The embodiment approach highlights the importance of studying AI variables, even in simulated and virtual agents, within a realistic systemic context to capture intelligence as an emergent property. Moreover, several robotics researchers believe that the investigation of AI requires a physical device such as an autonomous mobile robot to fulfil the criteria of being embodied and situated [66].

Another perspective to take into account in AI is the distinction between the creation of specialized intelligent artefacts and broader forms of intelligence general enough to solve problems of different types and under various conditions. These latter forms are representative of artificial general intelligence [67], an emergent trend within the so-called new AI, referring to a minimal cognitive system as found in complete autonomous agents such as mobile robots [68, 69], which put emphasis on learning functions.

This paper stands at the crossroads between AI, cognitive science, neurosciences and robotics in the context of the habituation phenomenon. Recent research has advanced the knowledge of the natural habituation process, inspiring us to emulate this complex learning function in order to study it and extend conclusions from a complete AI agent's perspective.

In particular, the multiple mechanisms behind the habituation mechanism suggest the possibility of dealing with different temporal sets of stimuli, producing adaptive behaviours in response. Studying the natural habituation process [70] can also support the hypothesis that biologically inspired modelling is a sound choice for building AI agents. In this regard, the impact of modelling a deeper bio-inspired habituation rule and its effects on the adaptive behaviour of AI agents is yet to be determined.

Beyond the design of a detailed habituation rule, the issue addressed in this paper concerns how an artificial agent should cope with the relatively huge amount of repetitive stimuli arising in various modalities and over different timescales in the real world. Nature found a partial response in the habituation process whose outcome is presumably to filter out redundant and persistent data. In fact, perception in cognitive agents is limited and the available physical stimuli far exceed their computational capacities. As a biological solution, the higher-level cognitive attention process permits to focus on some stimuli, but the exact mechanisms are still unknown. Which information should be ignored and how to handle it represent a real challenge for an AI implementation. Fortunately, the habituation process gives a clue to at least reduce the impact of repetitive stimuli that may be not necessary to get all the attention. Imitating this mechanism will enable AI agents to learn online how to ignore repetitive stimuli coming from the real world. Specifically, based on knowledge from neurosciences and consistent with empirical data, we present a habituation model tested on a complete bio-inspired cognitive system, and validated through the ensuing behavioural results of virtual and physical robots. From the inclusion of a few temporal learning features to the algorithm, the robots will deal with many frequency patterns of stimuli. We show that implementing such a habituation rule helps robots keep focus on currently desired tasks regardless of the temporal patterns of stimuli. To reach these conclusions, it seems natural to embed the learning rule in an ASNN as the core of the robot's controller, within a mostly bio-inspired paradigm, as opposed to other algorithms or hybrid mathematical solutions that could implement the habituation learning function with less biologically plausible scenarios.

In summary, we demonstrate first that it is possible to model a multi-featured habituation rule in an ASNN, and second that this learning algorithm leads to adaptive behavioural responses when embedded in a robot with several temporal contexts. Our results suggest that empowering an AI agent with this non-associative habituation function improves the range of attention and adaptability by capturing and ignoring redundant stimuli through experience. In order to demonstrate the habituation model, simulation experiments were conducted in a virtual 3D-world and physical platforms, where robots perceive repetitive stimuli of different frequencies, both when the robots are at rest and move freely. The habituation process is analysed and tracked at multiple levels, from synapses, neurons, neural network to global robotic behaviour observation. The results reflect the main characteristics of the natural habituation process:

- (1) Exponential decay of the behavioural response for a repeated neutral stimulus.



- (2) Recovery back to the natural response when the stimulus ceases.
- (3) Potentiation of habituation when protocols are repeated among different ISI.
- (4) Fast and strong habituation/recovery response for short ISI; slow and weak response for long ISI.
- (5) Short-term memory (STM) observed for short ISI and long-term memory (LTM) for longer ISI.
- (6) Fast and strong habituation for stimulus of lower amplitude and the inverse for higher amplitude.

In section 2, we describe the methodology and basis of the ASNN model used to embed the habituation rule, and we describe the external shape and internal neural architecture of the simulated robot. An incursion into a physical robot platform is also briefly presented. Our results are presented in section 3. Finally, a discussion and future researches are founded in section 4.

## 2. Methodological approaches

The literature describes several biological experiments from which to draw inspiration for highlighting the application and role of habituation in a model AI agent. A short list follows:

- (1) The defensive siphon and gill withdrawal reflex of *Aplysia* [5] is one source of inspiration when studying a habituation model in the context of robotic applications and artificial life simulations. Extensive experiments have been conducted with this primary organism, the neural circuit of which is greatly understood. Many direct (electrical stimulation, water jet, stick touch) and indirect (Petri-tap-vibration, air-puff) stimuli are used in the touch modality. Touch tolerance behaviour may be suitable for a robot to socially closely interact as a theoretical but plausible scenario.
- (2) The freeze and startle responses induced by flashes of light, tactile or sound stimulation in rats represent common protocols used to investigate the habituation process [9]. Sudden changes in the environment are known to draw the attention of biological organisms. Also, enhanced arousal behaviour seems to be correlated with the organisms' 'stop all activity' reflex. This behaviour has a protective side effect in biology, if we refer to the situation of a prey freezing so as not to be perceived by a predator, because of the sensibility gain in eye-brain movement over static frame observations. Experiments on the ability of toads [57] to distinguish between prey and lures focuses on habituation within the visual system at all neural relay levels. Enhancing attention drawn by novel stimuli by stopping current behaviours may also be wise for mobile robots.
- (3) The Petri dish tap stimulus and the touch reflex circuitry with *C. Elegans* [6] are also classical procedures where, during the habituation process, one can observe a gradual reduction of the normal backward swimming response when several temporal stimulus patterns are applied. Again, AI agents might also respond in this fashion as a precautionary behaviour to repetitive stimuli.

All of the previous scenarios capture the essential components of the habituation process at various levels of analysis. We propose to apply a similar strategy of investigation to a model AI agent, using AI-SIMCOG [71], a software package that offers an environment to develop and analyse virtual and embodied bio-inspired autonomous robots thanks to merging a 3D-world simulator [72], an ASNN editor and a transfer function for physical robot platforms. The characteristics of AI-SIMCOG's ASNN classify it in the leaky integrate-and-fire family, allowing the presentation of a deeper bio-inspired perspective of the habituation process, modelled as an integrative online synaptic property.

First, we show the features of the presented habituation rule by applying different repetitive temporal patterns of non-localized mechanical vibration stimuli to modulate, through experience, a backward movement as negative behavioural response of a simulated robot at rest. Then, a second round of simulations shows the impact of the habituation model when the robot moves freely.

The next section provides a description of the robot, followed by the details of its internal neural architecture. Next, we outline the expected robot behaviour, and then give a brief description of the dynamics of the ASNN with an emphasis on the habituation algorithm. Afterwards, we explain our experimental protocol. Lastly, we present two simulations with physical robots as proofs of concept before concluding the paper with discussions and perspectives for future work.

### 2.1. External structure of the simulated robot

The virtual robot (figure 2, left) is similar to common physical robot frames and includes an external circular structure and a two-wheel configuration to move about a simulated 3D-world. In this study, we focus on tactile, vibration and light stimuli which are relevant in both artificial and natural worlds. Therefore, the external design for the robot includes one vibration sensor in the centre and one light sensor disposed at the front. The vibration sensor perceives its modality over a 360° range and the visual sensor over a 60° cone of light ahead. In addition, four tactile sensors cover 90° each. These close-range sensors prevent dead-end situations when the robot makes contact with the walls delimiting the virtual world. All types of transducers return a single scalar value as a percentage of their field ranges.

### 2.2. Internal structure of the simulated robot: neurons, synapses and transducers

The neural architecture (figure 2, right) is composed of one transducer for each sensory neuron: one vibration, one light and four tactile. One intrinsic oscillatory neuron is added to generate dynamicity in the ASNN. Two motoneurons drive each wheel. One pair is excitatory and leads to a forward motion and the other is inhibitory inducing backward movements. Sensory neurons synapse onto motoneurons. The habituation rule is disengaged for the oscillatory neuron since it receives no input. The habituation function is also disengaged for the tactile neurons which just smooth the reflex responses of backward and turn movements, adding

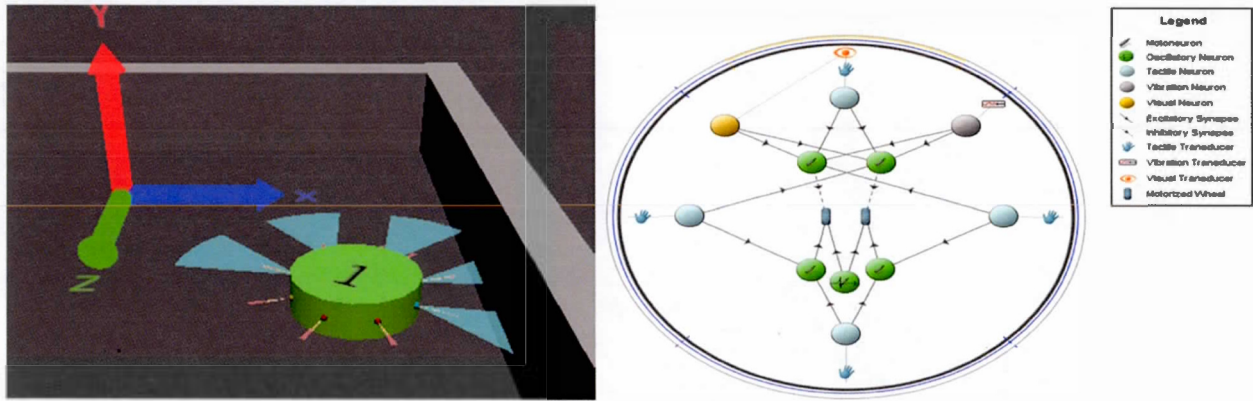


Figure 2. Left: example of robot's external structure and sensors ranges. Right: neural architecture.

Table 1. Initial synaptic weights.

Synaptic weight table		
Name	Strength	Type
S-Light-Mot-I-L	100	Excitatory
S-Light-Mot-I-R	100	Excitatory
S-Mot-E-L-LW	100	Excitatory
S-Mot-E-R-RW	100	Excitatory
S-Mot-I-L-LW	100	Inhibitory
S-Mot-I-R-RW	100	Inhibitory
S-Osc-LW	50	Excitatory
S-Osc-RW	50	Excitatory
S-Tact-E-Mot-E-R	80	Excitatory
S-Tact-E-Mot-I-L	90	Excitatory
S-Tact-N-Mot-I-L	100	Excitatory
S-Tact-N-Mot-I-R	75	Excitatory
S-Tact-S-Mot-E-L	100	Excitatory
S-Tact-S-Mot-E-R	75	Excitatory
S-Tact-W-Mot-E-L	85	Excitatory
S-Tact-W-Mot-I-R	100	Excitatory
S-Vib-Mot-I-L	100	Excitatory
S-Vib-Mot-I-R	100	Excitatory

unnecessary complexity to the simulation. Still, in a close world with the aim of increasing the frequency of encountering stimuli, these close range sensory neurons are helpful.

The design of the inter-neural type connections and the associated synaptic weights has a direct impact on the robot's behavioural possibilities. The start-up synaptic weights (tables 1 and 2) were intuitively initialized and adjusted to

produce basic locomotion behaviours (forward, backward, turn left or right) and avoiding stimuli (vibration, tactile, light). The neural network was set to contain a slight imbalance in the synaptic weights; this functional asymmetry helped prevent dead-end situations. With the exception of the inhibitory synaptic links from two motoneurons connecting to the wheels, all other neural links are excitatory.

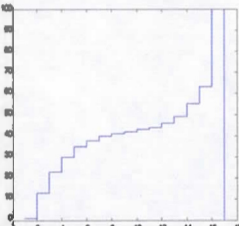
The vibration neuron is connected to the inhibitory motoneurons. Therefore, this type of stimulus will generate a momentary backward movement until it ceases when the robot is in a stationary state. If the robot is in a forward move state, a kind of transient freeze response is observed, depending on the motion speed and the stimulus strength. The light/visual neuron is also connected to the inhibitory motoneurons, leading to the similar backward behaviour when stimulated by this sensory modality. The four tactile neurons allow touch perception all over the structure of the robot. These sensory neurons connect to different pairs of motoneurons; the north tactile neuron is connected to the inhibitory motoneurons, the south tactile neuron is connected to the excitatory motoneurons, the east tactile neuron is connected to the left inhibitory and the right excitatory motoneurons, the west tactile neuron is connected to the left excitatory and the right inhibitory motoneurons. These chosen link patterns ensure avoiding prolonged contact with solid objects and justify their stronger synaptic weights and responses. Finally, the intrinsic oscillatory neuron is connected to the excitatory motoneurons for a constant pulse forward motion purpose.

Table 2. Matrix used for the numerical percentage scalar numbers of post-synaptic and sensory local potentials. Each entry corresponds to a specific temporal sequence of membrane potential variation.

0,0,0,0,1	0,0,0,1,1	0,0,0,2,1	0,0,0,2,2	0,0,1,1,1	0,0,1,2,1	0,0,2,2,1	0,0,2,2,2	0,1,1,1,1	0,1,2,1,1
0,1,2,2,1	0,2,2,2,1	0,2,2,2,2	0,2,3,2,1	0,2,3,2,2	0,2,3,3,2	1,2,2,2,1	1,3,2,2,1	1,3,2,2,2	1,3,3,2,2
1,3,3,3,2	2,3,3,3,2	2,4,3,3,1	2,4,3,3,2	2,4,3,3,3	2,4,4,3,1	2,4,4,3,2	2,4,4,3,3	2,3,5,3,1	2,3,5,3,2
2,3,5,3,3	2,4,5,3,1	2,4,5,3,2	2,4,5,3,3	2,4,5,4,2	2,4,5,4,3	2,4,5,4,4	3,5,4,3,2	3,5,4,3,3	3,5,4,4,2
3,5,4,4,3	3,5,4,4,4	3,5,5,4,2	3,5,5,4,3	3,5,5,4,4	3,5,5,5,4	3,6,5,4,3	3,6,5,4,4	3,6,5,5,4	4,6,5,4,3
4,6,5,4,4	4,6,5,5,3	4,6,5,5,4	4,6,6,5,4	4,6,6,5,5	4,7,6,4,4	4,7,6,5,4	4,7,6,6,4	4,7,6,6,5	4,7,6,6,6
5,7,6,6,4	5,7,6,6,5	5,7,7,5,5	5,7,7,6,5	5,7,7,6,6	5,7,7,7,6	6,7,7,7,6	6,8,7,6,5	6,8,7,6,6	6,8,7,7,5
6,8,7,7,6	6,8,8,6,6	6,8,8,7,6	6,8,8,7,7	6,8,8,8,7	6,9,8,6,6	6,9,8,7,6	6,9,8,7,7	6,9,8,8,7	6,9,9,6,6
6,9,9,7,6	6,9,9,7,7	6,9,9,8,7	6,9,9,8,8	6,9,9,9,8	7,9,8,7,6	7,9,8,7,7	7,9,9,7,7	7,9,9,8,7	7,9,9,8,8
7,9,9,9,8	8,9,8,7,7	8,9,8,8,7	8,9,8,8,8	8,9,9,8,7	8,9,9,8,8	8,9,9,9,7	8,9,9,9,8	8,9,9,9,9	9,9,9,9,9



**Table 3.** Left: neural parameters. Right: plot of the membrane potential.

Membrane potential variation: 1, 13, 23, 30, 35, 38, 40, 41, 42, 43, 44, 46, 49, 55, 63, 100	
Refractory period: 1 cycle	
Initial membrane potential: 41	
Resting potential: 41	
After hyperpolarization: 38	
Threshold potential: 63	

### 2.3. Expected robot behaviour

The robot's behavioural repertoire is defined by its external structural configuration, the neural network controller, the synaptic type connections, the associated synaptic weights, the learning functions used (habituation in this experiment) and the objects present in the virtual 3D-world. All behaviours are produced by a dynamic loop that entails sensing the virtual 3D-world, integrating the stimuli online through the ASNN and generating actions in the environment. The robot uses essentially a sensor-based feedback strategy to behave. Adaptation through learning experiences will be showed with the habituation process implemented at the synaptic level between the vibration and light sensory neurons and the motoneurons.

Robot locomotion depends on the produced temporal spike sequences combined with the synaptic weights of the links to the motorized wheels. Without any external stimuli, the robot moves straightforward because of the endogenous neural oscillator connected to the associated motoneurons and the motor wheels. This biologically plausible pacemaker generates periodic firings and the robot's displacement and speed are constant in the absence of obstacles or stimulations caught by its sensors. Whenever the robot encounters a wall, the tactile sensory neurons and the asymmetric weight connections produce a path change as a small movement backward and a slight turn to the opposite side. When the robot reaches a light or vibration zone, reflexive backward movements are observed. Habituation will modify these responses.

### 2.4. Dynamics of the ASNN

In this point, we provide details about the intrinsic dynamics of the ASNN model in AI-SIMCOG. First, the algorithm scans all transducers for stimuli inputs at each cycle. Depending on the field range, its mean intensity is calculated and converted to a single scalar percentage number. This value is multiplied by the efficiency (0–100%) of the transducer attached to its sensory neuron to obtain a new value. This value then serves as a pointer to a row vector within a scalar matrix. The vectors are meant to reflect the temporal amplitude variation of simulated current that contribute to the membrane potential. The matrix is hard-coded and enables the pairing of any response described by a vector component with all possible stimulus strengths (0–100%; see table 2).

The vector values follow an alpha-function shape [16, 42], nonlinearly summed up with all other inputs into the current neural state (refractory period, action potential or integrative

phase). The modulation of the membrane potential comes also from the applied learning functions (adaptation) and virtual ionic pumps that constantly push to maintain the homeostasis of the neural 'electrical charges' across membrane. This last phenomenon is not fully compensated by the intrinsic oscillatory neuron property which substantially leaks current at every cycle to periodically generate spikes. The neural integrative phase includes a threshold function that generates a spike to the connected components if the membrane potential crosses a certain trigger point. This is summarized in the next equation where  $v_m(k)$  is the membrane potential at cycle  $k$  and  $v_t(k)$  is the sum of transducer outputs:

$$v_m(k) = f_k \left( \sum v_t(k) + \dots \right) \quad \text{with} \quad f_k(x) = \begin{cases} 1 & \text{if } x \geq \theta \\ 0 & \text{otherwise} \end{cases}$$

The whole neural dynamics depend on the designer's choice of parameters. Table 3 provides the ones used in this work. A combination of ascendant exponential functions to reach a threshold value, an absolute refractory period, then a descendant logarithm function to a hyperpolarization state after a spike and finally another ascendant exponential function to reach a resting membrane potential compose the major elements to set. In the present design, the previous variables mentioned were all set for all neurons.

When a spike occurs, a similar calculation to that of the transduction phase is applied. Depending on the synaptic strength, a single scalar number (0–100%) is generated and identifies the vector corresponding to the variation (excitatory or inhibitory) of membrane potential to be received by the post-synaptic elements. If a learning function such as habituation is activated in the ASNN editor, a set of rules is applied (described below) to calculate a new percentage value. This modulated post-synaptic potential (PSP) moves the pointer in the matrix, producing a different variation sequence of the membrane potential in the post-synaptic element. The latter elements could be other neurons (including motoneurons) or directly effectors (muscle-like or motors). Finally, the sequence received by an effector leads to actions in the virtual or real world as dictated by another transform matrix that helps describe and understand the complete AI agent at the behavioural level. All these previous events are performed in batch at each step cycle. Thus, the ASNN in AI-SIMCOG refers to a leaky integrate-and-fire neural model with dynamical synapses and learning functions. Figure 3 summarizes the algorithm and ASNN functioning.

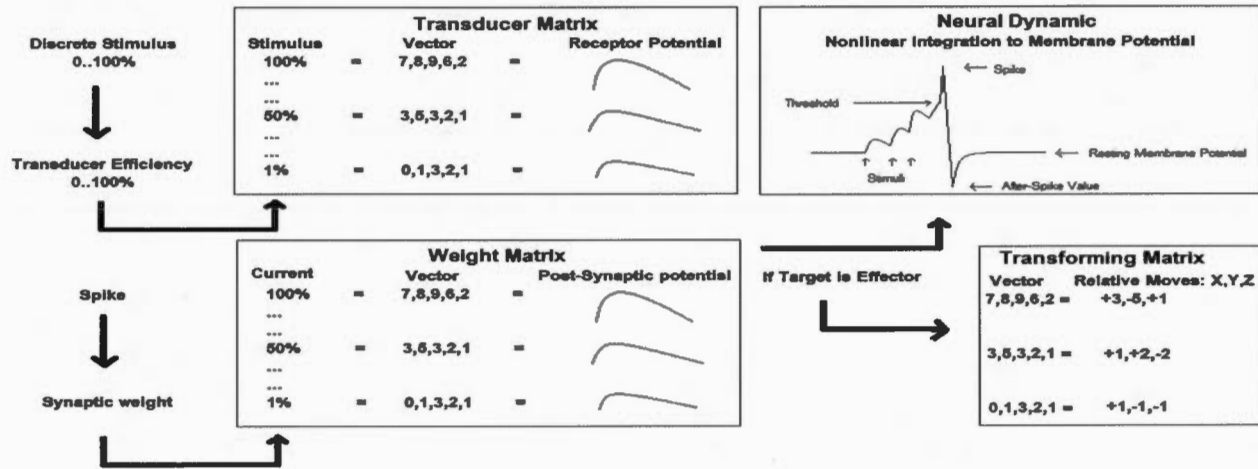


Figure 3. Logical algorithm steps of SIMCOG's ASNN.

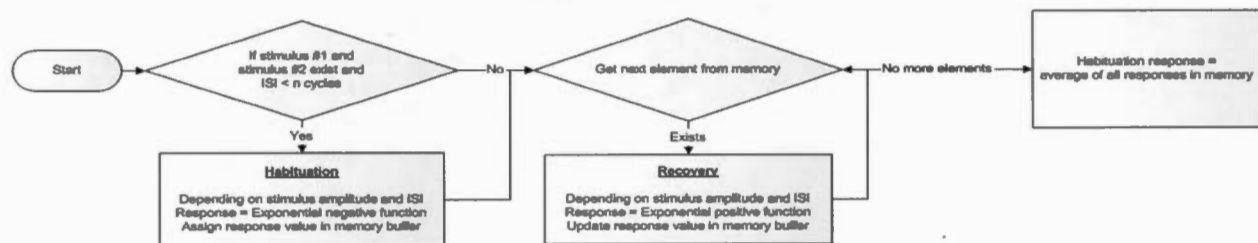


Figure 4. Habituation algorithm.

Table 4. Matrix of the percentage scalar numbers for the habituation rule.

ISI↓	Stimulus		Strength↓		
	0–5%	6–10%	11–25%	26–50%	51–100%
0–5 cycles	55,30,20,15,13	50,25,15,10,8	45,20,12,8,5	40,15,10,7,4	35,12,8,5,3
6–10 cycles	60,35,25,22,20	55,31,22,19,17	50,25,19,16,14	45,20,15,12,10	40,15,10,8,7
11–25 cycles	65,40,33,31,30	62,39,31,29,28	60,35,30,28,25	55,30,25,20,15	50,25,15,12,10
26–50 cycles	75,55,48,45,43	71,50,44,42,41	70,45,42,40,35	65,40,38,35,30	60,35,30,25,20
51–100 cycles	85,70,65,62,60	83,66,61,60,59	80,65,60,55,50	75,60,55,50,45	70,55,45,40,35

### 2.5. Habituation implementation

The habituation algorithm is based on the qualitative curve shapes described in the experimental literature [32]. The function consists of multiple conditional paths (figure 4) implemented in the ASNN controller of the complete AI agent. Using multiple pointers in the conditional tree, the habituation rule keeps track of the recent spike history in terms of ISIs for each synapse involved. Since the spikes history reflects the nonlinear integration of previous receptor potentials produced by the magnitude and ISI of external stimuli, this process results in particular sequences of spikes, with the correspondent temporal patterns subject to a habituation response at the synapse between the concerned pre- and post-synaptic neurons.

Specifically, the habituation rule produces a new coefficient value at each cycle (see table 4), constantly modulating the synaptic weights as a result. The coefficient depends on which pre-determined category the ISI history

falls into, after averaging all the launched sequential spikes within a specific temporal window as a result of the caught external stimuli. Consequently, merging all ISI patterns online leads to many outputs in the learning matrix, with a single coefficient at the end, a scalar percentage value dynamically affecting the synaptic weight. ISI categories were defined to correspond to five non-uniform lengths per unit cycle of the algorithm: 1–5, 6–10, 11–25, 26–50 and 51–100. The number of categories and their limits can be set otherwise as long as they allow simulating the general shape of the behavioural responses. Consistency with experimental parameters could also be adjusted in real time, but it was not a goal of this work to specifically match experimental data.

When different values of stimulus strength and ISI arise simultaneously in the same time window, an average is done. Because habituation is brought forth by a set of repetitive external stimuli, a recovery phase follows when the stimuli cease. As the recovery function is not simply the reverse of habituation, a different set of coefficients is used for the



backtrack sequence. If habituation occurs during the middle of a recovery phase, then an average of impact is computed. The same principle is applied in the case of multiple co-occurrences of habituation and recovery phases. Thus, both habituation and recovery coefficients move continuously across two vector tables from cycle to cycle, depending on the historical stimulus temporal sequence, including the stimulus magnitude, the ISIs and the current phases of habituations and recoveries.

## 2.6. Experimental protocol

This section shows the main features of the habituation process (decrementing response, recovery phases, potentiation, effect of multiple simultaneous ISI and bi-modalities). For each stimulus, we varied the amplitude, ISI and duration properties, and each block of experiments consisted of five consecutive trials to demonstrate the potentiation of habituation. The number of stimulus repetitions for each trial within a block was fixed to six to allow for the asymptotic part of the decaying exponential function to be present. For each stimulus, the normalized intensity was set to 10%, 25% and 50%, and the ISI was set to 10, 25 and 50 cycles of algorithm. The stimulus duration was set to 1, 3 and 5 cycles, and the time delay for the test phase after each trial varied from 10 to 100 cycles, preventing the complete reset of the habituation process. In a natural organism, the full recovery period is highly variable and depends on stimulus type, intensity and the input temporal pattern. In our experiment, we used a maximal recovery period of ten times the time it takes to reach full habituation for this parameter. For example, a stimulus that is totally habituated with an ISI of 100 cycles will fully recover after 1000 cycles. For all conditions, the robot completed the test with and without the habituation rule with the same initial parameters. Among the many scenarios to explore, we focused on three different situations permitting to highlight all the behavioural features of the habituation process:

- Block A: a wave stimulus of constant intensity is created at the arena's centre with a vibration range effect covering the entire world. The timing of the emission depended on set parameters.
- Block B: simultaneous vibration and light stimulus are perceived by different sensory neurons, and affect the same motoneuron to show bimodality integration through the habituation process. The vibrations and flashes of light are produced with the same temporal and amplitude parameters.
- Block C: three vibration waves' stimuli blink simultaneously with their own time-scale patterns.

The behaviours obtained stem primarily from the built-in neural circuitries, but adaptation comes from the plasticity provided by the non-associative habituation learning mechanism. Beyond demonstrating how the learning rule affects the weight variation observed in a static robot, the last part of the protocol shows the impact of habituation when a robot moves freely in a closed environment, but is perturbed by two different kinds of stimuli simultaneously, in different combinations of temporal and intensity patterns. Finally, the repeatability and sensitivity of the virtual experiments is 100%, thanks to the fixed initial conditions.

## 2.7. Experimentation with physical robots

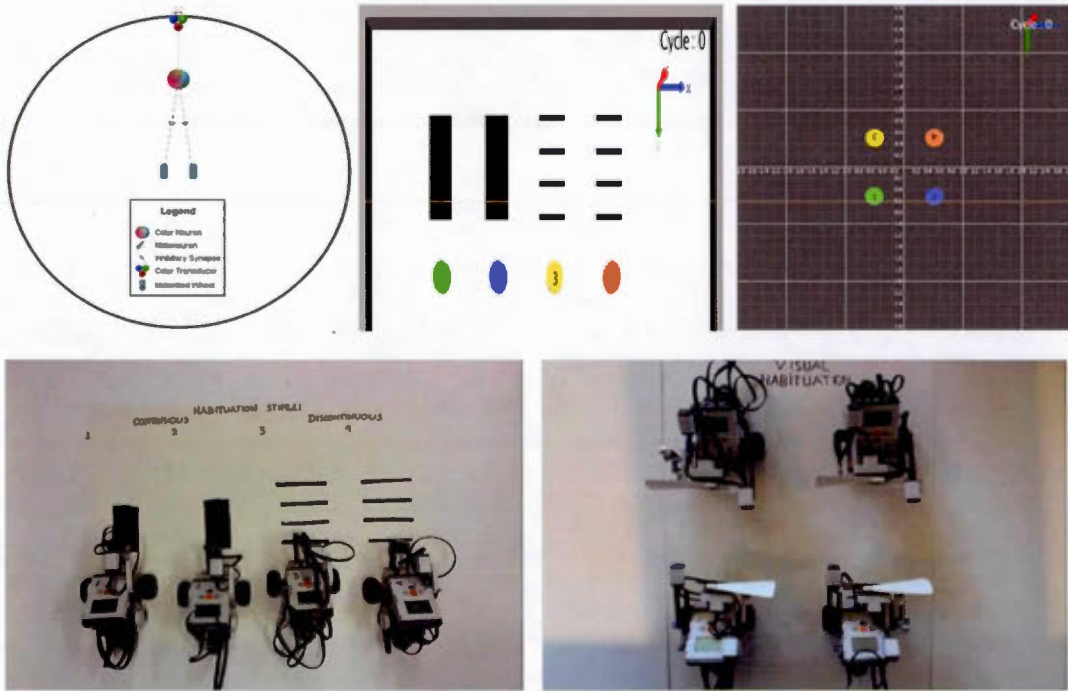
As a proof of concept, we went one step beyond AI virtual agent models by embedding the learning rule and the ASNN model into physical Legomindstorm NXT (Lego Co.) robots. In this part of the protocol, the neural architecture includes or lacks the habituation function (figure 5, top-left). It has one visual transducer responding maximally with the black colour (a bar on the floor) or the white colour (a structural component added to the robot). The colour transducer connects to its sensory-motoneuron, linking to the wheels with inhibitory synapses. We also added a constant forward motion to move the robots ahead.

Two simulations were first conducted in the virtual 3D world simulator (figure 5, top middle and right). Then, the same context was applied to physical robots evolving in the real world (figure 5 bottom left and right). The first simulation aimed to show how different stimuli patterns of black bars on the floor are habituating. The large black bars refer to stimuli of long duration in comparison to the tiny repetitive strips context. The second simulation demonstrated how nonlinear temporal stimuli patterns (different timing of the perceived white structural component in the approach of the opposing robot) impact the habituation mechanism. The repeatability and sensitivity of the physical experiments was not 100% because of the precision of the sensors/motors and the stimuli timing were not perfect. Nevertheless, without any adjustment of the transferred ASNN to the robots, stable results were obtained in a few consecutive trials.

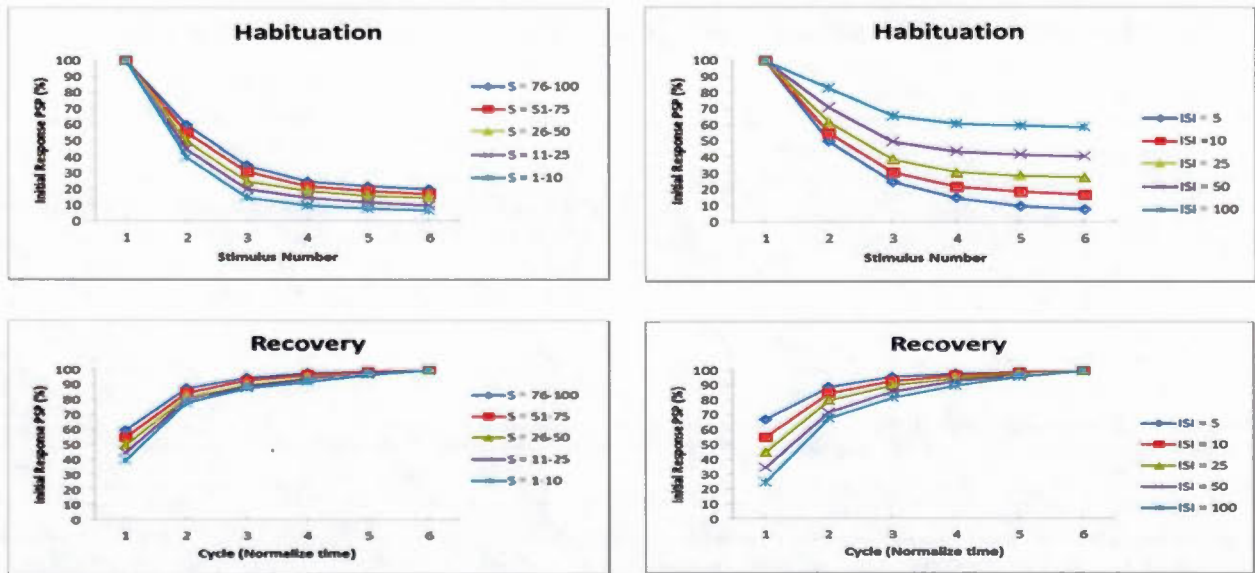
## 3. Results

Block A: figure 6 shows the presynaptic adjustment response for repetitive stimulus, plotted as an exponentially decreasing slope for the habituation process and an exponentially increasing slope for the recovery phase. Due to lack of space, only a few results are reported. First, for stimuli of variable strengths with one cycle duration and an ISI of ten cycles, we observe a maximum of habituation with the minimal strength (top-left). For stimuli strength of 10% and one cycle duration and variable ISIs, the shorter the ISI the faster and stronger the habituation response (top right), with a maximum effect for the lowest stimulus strength and the shortest ISI. At the bottom left and right of the figure, we observe the recovery response for the same parameters as at the top, showing a positive exponential function that is not symmetric to that of habituation, not only in the time domain, but also in the curve's general shape. Again, the recovery is stronger and faster for short ISI when a minimal strength stimulus is chosen (10% in this example).

In biological agents, the stimuli are perceived by transducers, and sensory receptor adaptation occurs if the stimulus period exceeds the transducer sensitivity. In our algorithm, the transducers' sensibilities are tuned to capture any events within a timescale of one cycle, which corresponds to the smallest time unit allowed to produce physical stimuli in the simulation. To determine the real effect of a duration stimulus from habituation, but without transducer adaptation,



**Figure 5.** Top-left: neural architecture of the robots. Top-middle and bottom-left: two different contexts of dark bars patterns on the floor, where robot1 and robot3 used the habituation rule and robot2 and robot4 lack the learning function. Top and bottom right: robots face up and are equipped white structural components to demonstrate visual habituation. Robot1 and robot3 include the habituation function, contrary as robot2 and robot4.

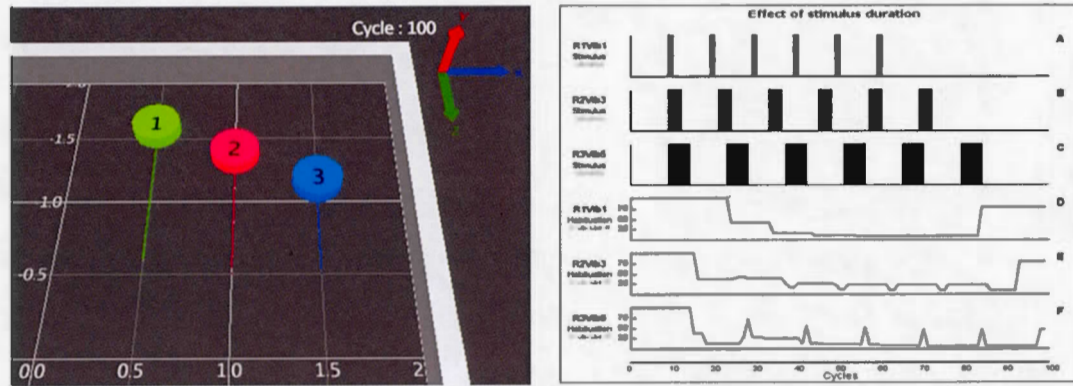


**Figure 6.** Top-left, habituation of response for different stimulus strengths with 1-cycle duration and 10-cycle ISI repeated 5 times. Bottom-left, recovery over normalized initial response time for different stimulus strengths with 1-cycle duration and 10-cycle ISI repeated 5 times. Top-right, habituation of response at different ISI for a stimulus of 10% strength and 1 cycle duration repeated 5 times. Bottom-right, recovery over normalized initial response time at different ISI for a stimulus strength of 10% and 1 cycle duration repeated 5 times.

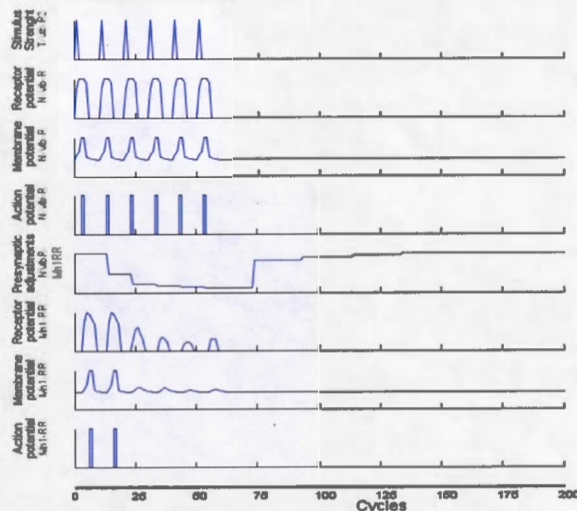
we should implement a lower transducer sensibility or produce stimuli at a higher frequency. In our simulations, habituation is applied without a transducer adaptation phenomenon. The result is that the longer the stimulus duration is, the more it

will raise the membrane potential fast and ultimately affect spike production, reaching a saturation ceiling of one spike every two cycles (500 Hz if one cycle corresponds to one millisecond) because of the minimal refractory period of one





**Figure 7.** Effect of stimulus duration (vibration in this case): the stimulus length is 1 cycle for Robot1, 3 cycles for Robot2 and 5 cycles for Robot3 for 6 repeats and an ISI of 9 cycles. On the left side, the robots path courses are illustrated. On the right side, the timing and different duration of the stimuli (A) to (C) are reported in relation of the correspondent presynaptic modulations from the applied habituation rule (D) to (F).



**Figure 8.** Demonstration of the habituation sequence.

cycle. Correlation with the ISI pattern also changes with the exact duration of the perceived stimuli, affecting the presynaptic parameter and leading to a more rapid habituation for the longest duration (figure 7).

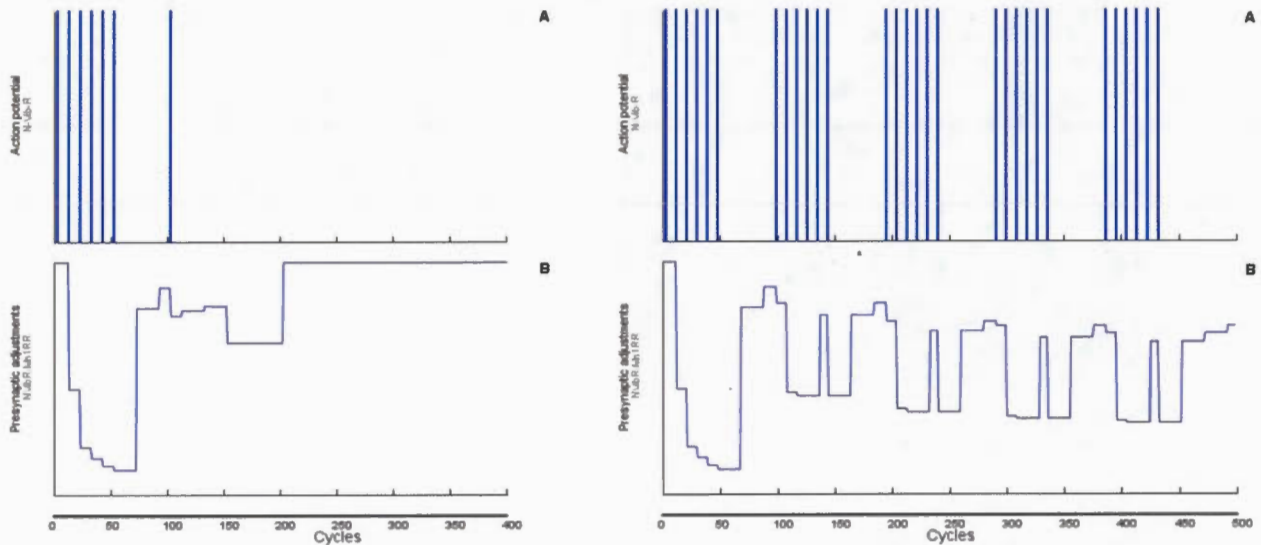
Here, the habituation process is demonstrated in a temporal event sequence where six vibration stimuli of 50% strength and nine cycle ISI, for one cycle duration each, are launched during the simulation. In figure 8, the stimulus is perceived by a transducer (T-Vib-R2) (A); then, it is converted into a local variation of receptor potential (B) that affects the overall membrane potential of the sensory neuron (N-Vib-R) (C); finally, a spike is produced (D). This occurs for each stimulus and, following the produced spikes, the habituation process takes place at the synapse between the sensory neuron and the motoneuron. As expected, a pre-synaptic adjustment is observed (E) because of the consecutive stimuli and the habituation function application, reflecting a down modulation efficiency of the incoming spike train. It can be seen that habituation happens when two or more consecutive stimuli

occur in a certain timeframe and the time delay between them is shorter than the internal value parameter that triggers restoration of the natural response or the recovery phase start time. Then, the received spikes are converted to a variation of receptor potential at the synapse (F), reflecting a down regulation as well as the induction of a variation of the membrane potential (G); consequently, the spike emission pattern induced by the motoneurons also follows a negative modulation (H).

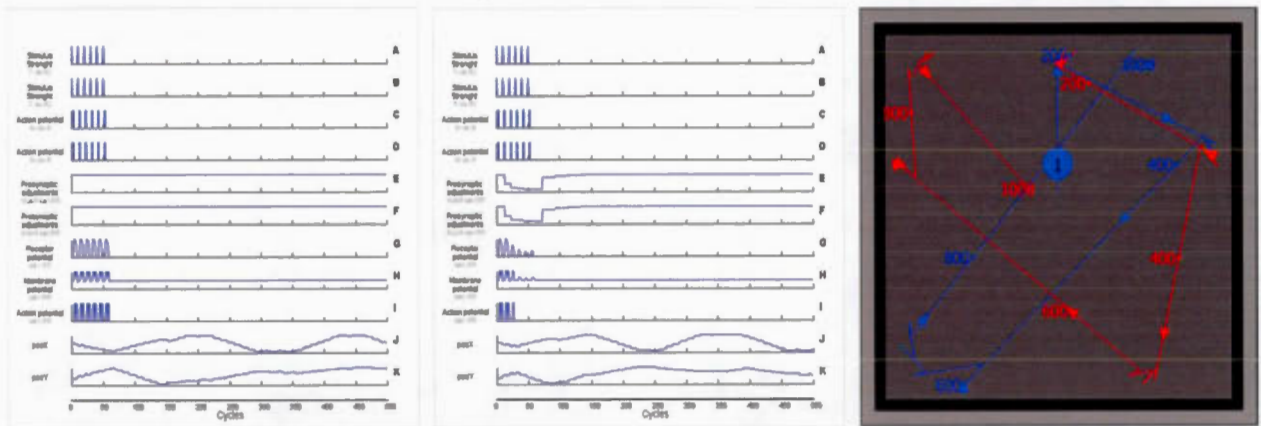
From this simulation, it can be seen that only two spikes were produced by the motoneuron because of the habituation process, instead of the six spikes expected from the six stimuli generated by the sensory-neuron onto the synaptic site of the motoneuron. As a result, the backward movement is less pronounced in the case of a robot implemented with the habituation function. In the comparative situation of an animal without habituation (not shown), the response to a repetitive vibration stimulus corresponds to a fixed, identical and reactive short backward movement, resulting in a reduction of the absolute distance travelled forward during the simulation. In one sense, the habituation function helps the animal to ignore a repetitive and neutral stimulus that could slow it down in its path course.

The habituation process can be potentiated in time by specific temporal patterns of events. To observe this short-term memory, the time delay between blocks is set just below the total recovery time of the synapse or before the last steps of the algorithm during the recovery phase. If a test stimulus is applied just before total recovery (figure 9(A), left), the synaptic efficiency parameter will keep track of the previous events generated in the anterior blocks. The synapse keeps a memory trace of the habituation while adjusting back into the recovery algorithm and not fully recovering (ISI = 50) (figure 9(B), left). The habituation process (figure 9(B), right) also shows the potentiation effect when multiple blocks in sets of repetitive stimuli occur close together (figure 9(A), right). The synaptic efficiency is progressively reduced in this protocol, due to the repetitive blocks of stimuli applied close enough in time to avoid complete restoration of the normal synaptic response. In summary, all the main characteristics





**Figure 9.** Short-term memory is shown (left) with a stimulus test occurring within 100 cycles, before full recovery. Potentiation is demonstrated (right) with the some steps back in the recovery algorithm.



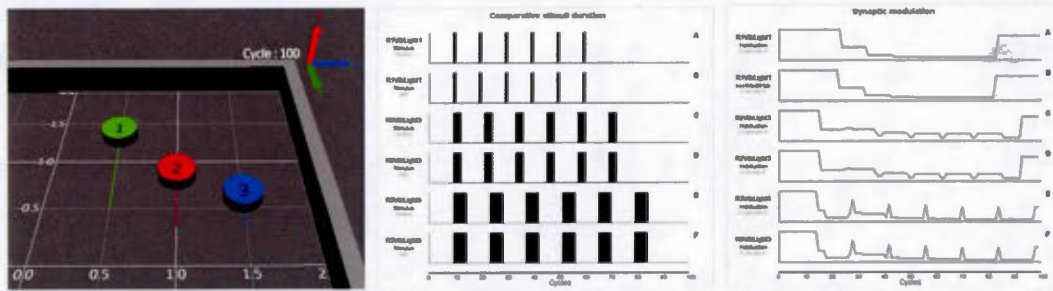
**Figure 10.** Left: bimodality integration without habituation. Middle: bimodality integration with habituation. Right: difference in the robot's path course. Blue trace is without habituation and red trace is with habituation.

of the habituation process mentioned in the introduction are demonstrated in this first part of the experiment.

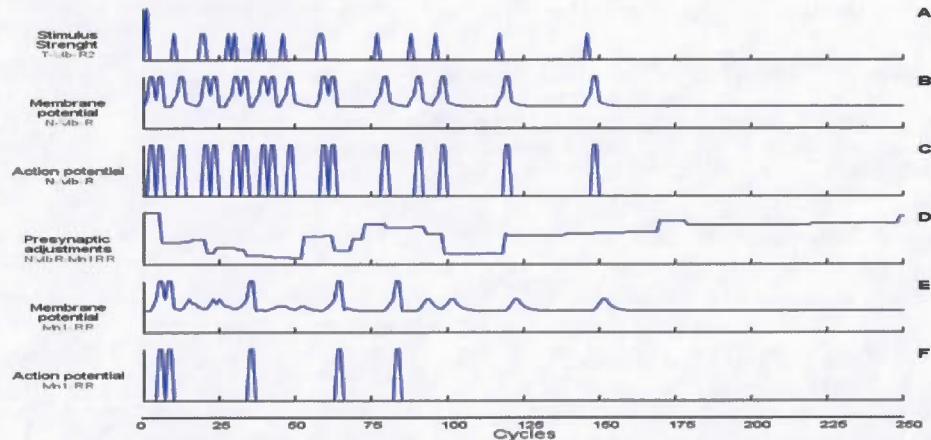
Block B: merging two stimuli of different modalities in one sensory neuron is the same as integrating two stimuli of the same modality for this neuron, reflecting the universality of electrical current fluctuation. Hence, one should expect the production of double amount of the excitatory post-synaptic potentials (EPSP) generated by spikes, reflecting stimuli of similar strength and temporal shape, and received at two different synaptic sites pacing at the same rhythm. Therefore, the habituation should happen in both synapses as demonstrated in Block A, according to the specific temporal pattern and strength features of the repetitive spikes. This means that for subsequent stimuli, all repetitive beats start habituated, leading to a negative exponential curve of the neural response. Then, adaptation happens at the presynaptic site between the sensory-neuron and the linked motoneuron, modulating the EPSP strength. Even if habituation slows

down stimulus effect, merging two stimuli provokes an acceleration of the spikes generated from the motoneuron to the motor in comparison to a one-synapse receiver setup. However, the electrical activity is not doubled because of the nonlinear integration of the membrane potential. To illustrate the bimodality effect, we chose a sequential graph (figure 10) and drew the animat paths for 1000 cycles.

Figure 10 (left and middle) shows stimuli that occur in total temporal coordination at the transducer site for both modalities (A and B). Also, they trigger spikes in the same manner within their respective sensory neurons (C and D). In figure 10 left, the habituation function is turned off at both synapses (E and F); in figure 10 (middle), synaptic adjustments due to the habituation function have been made. Interestingly, after spike reception at the motoneuron, we already observe a change of receptor potential between the two simulated robots (G) as well as in the membrane potential variation (H). As expected, the spike pattern is different depending on the



**Figure 11.** Left: path traces for bimodality when combine with a stimulus duration effect. Middle: comparative graphic showing the difference in stimuli duration in both modalities. Right: differences at the presynaptic site.



**Figure 12.** Temporal sequence of three sources of vibration input occurring in different phases, perceived by one transducer and leading to a habituated stimulus.

state of the learning function (I). Finally, the effects on the motorized wheels, the movement behaviours of the animat and its absolute position in the environment (J and K) are quite different during the first 200 cycle period, reflecting the effect of the habituation and recovery phases. More precisely, the backward movement is reduced because of the habituation function, altering the animat's path course. Path traces are different depending if the habituation rule is turn on (red) or off (blue).

In a robustness test of the habituation rule, we combine (figure 11) the bi-modality and stimulus duration effects (initial conditions of figure 7), showing differences in the robot's path course. When comparing to the results of both figures, this difference happens not to be additive, but is nonlinearly integrated.

Block C: when multiple sources simultaneously emit signals out of phase, tracking the individual ISI patterns is complex, and the overall temporal differences dictate the behaviour of the presynaptic modifications. To differentiate the timelines of each event, many transducers and neurons should exhibit their own signature for a specific bandwidth or a desired phase at which to capture or emit stimuli. These mechanisms support biological data and imply a heterogeneous neural network with some of the neural units showing the properties of specific filters, frequency tuned or input selective, mainly in the temporal domain. This is definitely beyond the capacities of the present ASN algorithm model.

In this experimental block, three emitting sources of vibration generate stimuli with different ISI patterns, starting at the second cycle with 50% strength and six repetitions, for a total simulation duration of 500 cycles. In the example below, the ISI for the three different sources was set to 9, 19 and 29, leading to the following temporal spike patterns (each value represents the cycle number of the time-event):

*Source 1:* 2, 11, 20, 29, 38, 47

*Source 2:* 2, 21, 40, 59, 78, 97

*Source 3:* 2, 31, 60, 89, 118, 147.

A transducer with enough sensitivity to capture all stimuli within one cycle will result in the three independent stimulus sequences shown above being reorganized into the following linear sequence of events (see also figure 12(A) below):

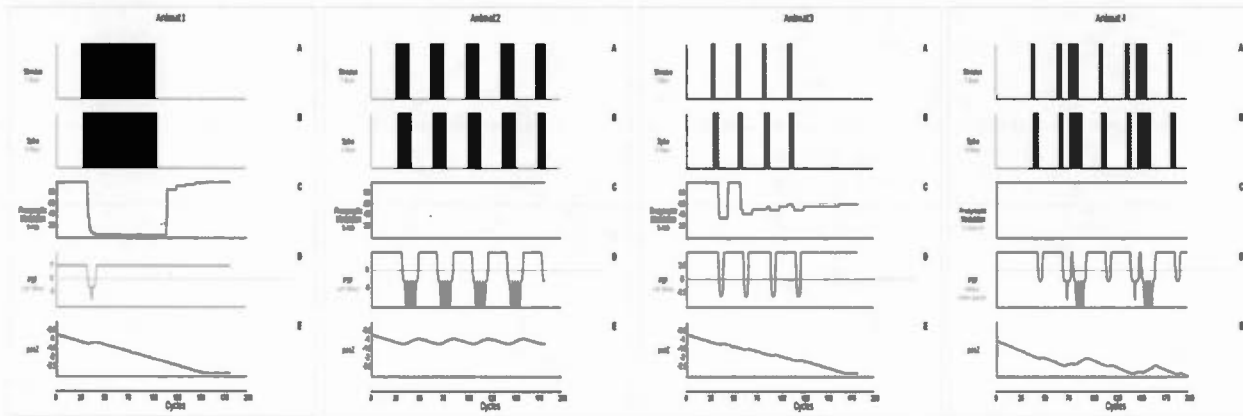
*Source 1 to 3:* 2, 11, 20, 21, 29, 31, 38, 40, 47, 59, 60, 78, 89, 97, 118, 147.

From these stimuli, the membrane potential change of the sensory neuron (figure 12(B)) leads to an emission of spikes (figure 12(C)) that exhibit the same frequency as the simulated physical vibration stimuli. Thus, instead of perceiving three different sources beating individually at 9, 19 and 29 ISI, the actual perception at the synapse between the sensory-neuron and the motoneuron will be of one source exhibiting the following heterogeneous ISI pattern:

*ISI:* 9, 9, 1, 8, 2, 7, 2, 7, 12, 1, 18, 11, 8, 21, 29.

The spikes emitted by the sensory-neuron affect the membrane potential of the motoneuron (figure 12(D)).





**Figure 13.** (see also figure 4) Animat1 represents the robot with a habituation rule when facing a large dark bar on the floor (A) and the spike production of the sensory-neuron. It results in fast habituation (C) shown by a brief backward movement (D) and a relatively constant forward movement (E). In the same context, Animat2 lacks the function and, consequently, always avoids the bar by stepping back (D), (E) in a reactive behaviour. Animat3 is equipped with the learning rule and adapts progressively (C) to four black bars (A), while Animat4 slowly moves forward (E) because of its inner drive and reactive backward behaviour (D) when facing a black bar.

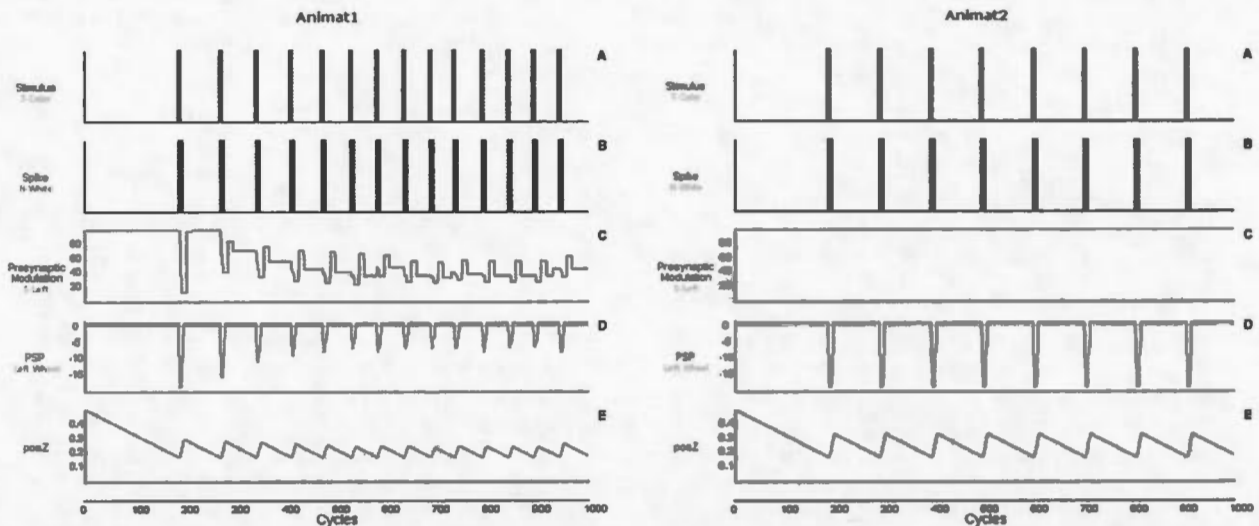
Presynaptic adjustments occur according to the temporal input sequence from the sensory-neuron through the synaptic component shared by the receiving motoneuron (figure 12(E)). The observed line shows an oscillation between multiple habituation and recovery phases overlapping in different timelines, leading to the observed output pattern from the motoneuron (figure 12(F)). In total, a mild adaptation occurs and fewer backward movements are produced, from which we can conclude that habituation still happens in a complex temporal pattern of inputs, even if repetitive stimuli came from multiple out-of-phase sources.

In this example, the response of repeated stimuli does habituate rapidly and strongly because of the small ISI value used, but partial recovery occurs when the ISI increases. Oscillations between phases of habituation and recovery are observed until stabilization after the final stimulus. The goal of this virtual experiment, fully reproducible since the initial condition is fixed, was to study the perception of and habituation to repetitive stimuli that do not follow a regular beat. Because of the transducer tuning limit, instead of responding to three different captured inputs, the algorithm responds with a middle strategy. The time difference between two stimuli leading to habituation will follow specific individual rules of recovery depending on the starting ISI. If another stimulus occurs in a different timeline than the previously habituated stimulus, a mean is computed for both the habituation and recovery phases. As another example, two repetitive stimuli of 9 ISI will generate approximately 60% negative adjustment of the natural response. If another stimulus occurs 30 cycles later, then the algorithm will combine the adjustment of the previous 9 ISI and 60% adjustment with the new 21 ISI. To summarize, any linear sequence of stimuli will be interpreted by the transducer via the mean of all previous habituated and recovered events, within the upper limits fixed in the temporal buffer memory. In natural models, biochemical cascade events are probably launched each time a stimulus is detected. This led us to also implement a trade-off between all possible parallel streams occurring in the process.

Finally, two simulations are related to physical experiments (see also figure 5). In the first simulation (figure 13), the robots move forward and perceive black bars on the floor. Depending on the inclusion of the habituation rule and the stimuli pattern, the behavioural responses change totally. Animat1 rapidly ignores the constant stimulus with the habituation function, moving forward pretty much. Animat2 which lacks the rule cannot move forward through the constant black bar stimulus and keeps avoiding it. Animat3 faces four tiny black bars on the floor and, with the rule progressively learn to ignore the stimuli. Animat4 (lacking the rule) moves backward in a reactive mode in front of the stimuli, but still move forward slowly because the strength of the constant forward force exceeding the negative backward movement provoke by the tiny black bars on the floor. In the second simulation, we show with a neural sequence graphic how a habituation learning rule (figure 14, left) could result in the appearance of social tolerance, by decreasing the response of avoiding white robots, while the situation is only reactive with the robots lacking the non-associative function (figure 14, right).

#### 4. Discussions and future works

In the preceding sections, we showed how the natural habituation process can be emulated in a virtual robot controlled by an ASNN. The process was implemented to allow complete bottom-up explanations, from stimulus to robot behaviours by using the AI-SIMCOG software which permits simultaneous step-by-step tracking of the process at many levels of scrutiny. The shapes of the obtained data curves show a good approximation of the major characteristics of habituation, although accurate biological modelling was not a priority in our work. The results reveal behavioural adaptation at the agent level, obtained by including this primary non-associative learning process in the ASNN. The major adaptive effect was to prevent unnecessary responses to repetitive stimuli. Extending the temporal features of the habituation rule



**Figure 14.** On the left, Animat1 and Animat3 behave similarly, slowly habituating white stimuli pairs (A) with spikes production (B), resulting in getting closer to each other in relation to the presynaptic modulation (C). The situation for Animat2 and 4 which both lack the learning function is to react with a constant behaviour of avoiding white object, resulting in wide backward movements (E).

empowers the robot to behave differently regarding the capture of repetitive stimulus patterns. Although our conclusions in this regard are positive, further discussion is warranted about the model's limits and power.

In experimental Block A, we demonstrated six essential features of natural habituation with the same computational rule implemented within an ASNN. Exponential decay to repeated stimuli and the following exponential positive slope for the recovery phase when stimuli cease were the minimal phenomena to model. The potentiation of habituation between block of trials was another concept allowing extension of learning in the temporal domain. All these temporal features were object of modulation in terms of response latencies and amplitudes as a result of the length of the timing between each stimuli event, whether it was between stimuli (ISI) or between block of trials. The intensity of the stimuli also offered a modulated response from the habituation function in our proposed model.

One interesting question that we addressed was whether habituation can work across a full temporal continuum or, more naturally, in finite steps related to the specific time constants of several possibly overlapping chemical cascade pathways. What is the effect of repetitive but heterogeneous stimuli patterns on learning and memory processes? In a natural agent, the behavioural implications raise questions about the choice of training protocols used and the type of organism involved. Abbott and Kandel [73] reverse the question by finding the optimal training set to optimize the behavioural response, further revealing the dynamics of the underlying process. The authors find that sets of the nonlinear timing of stimuli pattern seem to be the optimal solution for training, reflecting the dynamical complexity of the underlying biochemical processes. At the light of these findings, we may also infer that temporal randomization and symmetry of the caught information could also have a positive effect on the learning curve from the habituation mechanism.

Although our current model does not include all these findings, at least the habituation algorithm supports variable observed response to nonlinear timing of any stimuli patterns.

Another point to mention is that we deliberately avoided the use of fitness functions with genetic algorithms or other optimization methods for parameter tuning. This might have led to more efficiency of the obtained results. But, one can also argue that the habituation is a learning phenomenon that happens relatively quickly during the 'life' of an organism; this is not coherent with the typical evolutionary process optimization. On the other hand, it was relatively easy to produce a heuristic choice of percentage parameters and our computational model of habituation showed that repetitive stimuli lead to a temporally decreasing behavioural response due to modulation of the synaptic efficiency.

Even though behavioural predictions are difficult to establish in a multi-level description of a dynamic cognitive system, complex neural circuit architectures including second-order interneurons, polysynaptic contacts, central pattern generators and large neural network with a habituation rule should be investigated, to further elucidate the limits and power of adaptive behaviours production in robotic agents.

Proposing that the purpose of the habituation process is to capture novelties, emphasize selective attention and filter information redundancy is not new, but many aspects of it deserve scrutiny in the context of robotic implementations. Could habituation participate to other higher cognitive concepts as intuition or feelings? How often is this mechanism used in an organism's day-to-day existence? Does repetitive stimulation in a disordered or chaotic temporal manner always trigger habituation? What happens when the input is complex and comes from cross-modalities? What is its relevance in lifelong adaptation considering the variety of lifespans in natural organisms and the mostly fixed dynamics of the neural components involved in the habituation process? Above all, should AI agents, such as mobile robots, include this kind



of natural learning process? We designed blocks B and C of the experimental protocol in an attempt to start answering these questions. The situations that we investigated were very simple, but they shed some light on the possible implications of the habituation process on larger cognitive scenarios, paving the way to include higher natural complexity in future AI agents, both simulated and physical models.

In the bimodality test (Block B), two different sensory neurons generated spikes on a single motoneuron in the context of repetitive incoming stimuli from two different sources. This setup inevitably reduces the overall spike firing rate and leads to a decrease in the observed behavioural displacements. The present mechanism implies a negative down modulation applied at the presynaptic level. It results in the reduction of the post-synaptic potential, slowing down the motoneuron spike rate and reducing the effector activities through the nonlinear combination of the two perceived sources. One can easily imagine what happens to firing patterns when multiple sensors are repetitively stimulated, converging and merging onto a particular cell: the response should increase but not proportionally until it reaches saturation or the maximum firing limit controlled by the refractory period. Thus, habituation is just one more way of slowing down the input traffic at a specific synaptic site. In fact, for a natural neural cell, the incoming spikes from several different input sources are integrated indifferently, providing only spikes as output responses. Regarding the habituation process and the bimodality context as input sources, the neural architecture and connections involved should have almost equal impact on a neural cell's output behaviour.

Input patterns occurring simultaneously in several different physical modalities and at multiple temporal scales are the rule in the real world. Some sequences are probably distinct and well organized in the temporal domain, and different sources may be perceived at the same time by natural organisms. In Block C, we tried to figure out how to discriminate different temporal inputs at the level of an artificial synapse through the habituation mechanism. The idea was to reproduce different biochemical cascades involved in the natural process by dynamically switching the input sequence according to the chosen categories, based on their own temporal characteristics. This is a hypothesis that might create an opportunity to enhance the learning arsenal of AI agents via a more complex habituation function.

Synapses perturbed by a repetitive input adjust their response in a given direction but also return to their baseline level eventually when the stimuli cease. Keeping track of one repetitive stimulus pattern was easy because it was associated with a specific ISI and recovery category in the algorithm. When several stimulus patterns are captured simultaneously on several time-scales and different amplitudes, the online synaptic efficiency modulation turn out to be fairly difficult to anticipate. The perception of complex temporal input patterns may need other learning processes to fully grab the picture with an ASNN. This may be based on selective stimulus capture, as in the natural neural stimuli preference (i.e. frequencies in auditory cell and orientation in place cell) properties known phenomenon. Otherwise, achieving the goal of input demultiplexing computation could be obtained with recurrent

links from neurons to other synapses to maintain some signals. Also, neuron parameters such as the refractory period could be tuned to a specific value to filter a small bandwidth, so as to respond only to a precise temporal pattern. Moreover, different spatial synaptic configurations could capture time events differently; the sensor topology and the shape of the dendritic tree could also help resolve time-related problems. Finally, elaborating or dedicating functional artificial neural circuits such as the hippocampus could certainly compute data incoming from different sources in several temporal scales. Because understanding the world partially means perceiving situations in the time domain in order to behave and react appropriately, managing complex temporal scales of repetitive input pattern stimuli is a minimal capability to implement in an artificial entity. This was our motivation for exploring these situations, extending the habituation model in the time domain to respond to these evidences, based on current experimental knowledge but applied to artificial agents.

Habituation is involved in broader types of learning and memory, and it is only one piece in the puzzle of cognition and adaptation possibilities. Moreover, it may interfere or cooperate with other types of processes as in the spike-timing-dependent plasticity [74] learning rule. Although computational modelling is often a reductionist approach, it allows one to rapidly test hypotheses and explore the limits of a model's variables. However, simulations may not necessarily reflect the real world and the interpretation of the habituation algorithm that is presented in this paper may be risky without such realism. In fact, mimicking a natural process with a computational model in a multi-level system such as a small, dynamic artificial entity is full of bias inherent to the virtual setup.

Modelling biological properties for AI agent controllers in an attempt to simulate observations of the natural intelligence phenomenon is one way to proceed. In the fifties, the trend was based on logical propositions; later the connectionist movement put forward an alternative, an implementation of which being the ASNN. ASNNs are more complex, more biologically plausible, and also more able to account for the temporality aspects of cognition. One might ask: Do we really need to draw from such a deep biological inspiration in AI models? The black box between input and output in the real world has to be filled somehow and certainly other models will continue to be developed with other paradigms. Our position is that the best example of intelligence actually lies with living creatures. Natural organisms, however primitive they may be, exhibit frank signs of intelligence. As a bio-inspired robotics branch, we think that capturing the functional aspects of some chosen biological mechanisms and transferring those incrementally in an AI agent model represents a reasonable approach to reproduce the intelligence phenomenon. One of these mechanisms we tried to highlight in this article was the habituation learning process. Designing a small artificial neural circuit enriched with a low level but complex learning rule such as habituation revealed that the bio-inspired paradigm used in AI and robotic sciences could find a link at the behavioural level.

We also validated results with physical robots (section 2.5, see figure 5) and videos are available on the following web



site: [www.aifuture.com/neurocode/id/18/learning-functions/non-associative-habituation](http://www.aifuture.com/neurocode/id/18/learning-functions/non-associative-habituation). In these simulations, we can see how robots equipped with the habituation function gradually ignore dark strip stimuli on the floor or tolerate more the facing opposing robot.

In conclusion, our study showed that a habituation function can be included in an ASNN model. An ASNN with augmented learning content was embedded in an autonomous simulated robot behaving within a dynamic virtual 3D-world. This small cognitive system allowed us to simultaneously investigate the habituation process at many levels, from the description of presynaptic modulation and spike production to the observation of actual robot behaviour. Our intention was not just to include a standard habituation function in an ASNN algorithm; rather, we expanded the standard model to include several important recent known features of the natural process, grounding the computational model in a deep bio-inspiration to drive the AI agent's adaptive behaviour. In our model, habituation results in real changes based on experience; it helps robots to ignore various kinds of repetitive stimuli patterns and potentially concentrate on their main tasks.

## References

- [1] Thorpe W H 1956 *Learning and Instinct in Animals* (Cambridge: Harvard University Press)
- [2] Thompson R F and Spencer W A 1966 Habituation: a model phenomenon for the study of neuronal substrates of behavior *Psychol. Rev.* **73** 16–43
- [3] Groves P M and Thompson R F 1970 Habituation: a dual-process theory *Psychol. Rev.* **77** 419–50
- [4] Rankin C H *et al* 2009 Habituation revisited: an updated and revised description of the behavioral characteristics of habituation *Neurobiol. Learn. Mem.* **92** 135–8
- [5] Castellucci V F and Kandel E R 1974 A quantal analysis of the synaptic depression underlying habituation of the gill-withdrawal reflex in *Aplysia* *Proc. Natl Acad. Sci. USA* **71** 5004–8
- [6] Wicks S R and Rankin C H 1995 Integration of mechanosensory stimuli in *Caenorhabditis elegans* *J. Neurosci.* **15** 2434–44
- [7] Acevedo S F, Froudarakis E I, Kanellopoulos A and Skoulakis E MC 2007 Protection from premature habituation requires functional mushroom bodies in *drosophila* *Learn. Mem.* **14** 376–84
- [8] Rankin C H, Beck C D O and Chiba C M 1990 *Caenorhabditis elegans*: a new model system for the study of learning and memory *Behav. Brain Res.* **37** 89–92
- [9] Leussis M P and Bolivar V J 2006 Habituation in rodents: a review of behavior, neurobiology, and genetics *Neurosci. Biobehav. Rev.* **30** 1045–64
- [10] Sirois S and Mareschal D 2004 An interacting systems model in infant habituation *J. Cogn. Neurosci.* **16** 1352–62
- [11] Antonov I, Kandel E R and Hawkins R D 1999 The contribution of facilitation of monosynaptic PSPs to dishabituation and sensitization of the *Aplysia* siphon withdrawal reflex *J. Neurosci.* **19** 10438–50
- [12] Balaban P M 2002 Cellular mechanisms of behavioral plasticity in terrestrial snail *Neurosci. Biobehav. Rev.* **26** 597–630
- [13] Glanzman D L 2006 The cellular mechanisms of learning in *aplysia*: of blind men and elephants *Biol. Bull.* **210** 271–9
- [14] Hawkins R D, Kandel E R and Bailey C H 2006 Molecular mechanisms of memory storage in *Aplysia* *Biol. Bull.* **210** 174–91
- [15] Ye H-Y, Ye B-P and Wang D-Y 2008 Molecular control of memory in nematode *Caenorhabditis elegans* *Neurosci. Bull.* **24** 49–55
- [16] Rudy J W 2008 *The Neurobiology of Learning and Memory* (Sunderland, MA: Sinauer Associates)
- [17] Whitlow J W 1975 Short term memory in habituation and dishabituation *J. Exp. Psychol.* **104** 189–206
- [18] Vogel E H and Wagner A R 2005 Stimulus specificity in the habituation of the startle response in the rat *Physiol. Behav.* **86** 516–25
- [19] Burrell B D and Sahley C L 1998 Generalization of habituation and intrinsic sensitization in the leech *Learn. Mem.* **5** 405–19
- [20] Wang D 1994 Modeling neural mechanisms of vertebrate habituation: locus specificity and pattern discrimination *J. Comput. Neurosci.* **1** 285–99
- [21] Del Rosal E, Alonso L and Moreno R 2006 Simulation of habituation to simple and multiple stimuli *Behav. Proc.* **73** 272–7
- [22] Rankin C H and Carew T J 1987 Development of learning and memory in *Aplysia*: II. Habituation and dishabituation *J. Neurosci.* **7** 133–43
- [23] Beck C D and Rankin C H 1995 Heat shock disrupts long-term memory consolidation in *Caenorhabditis elegans* *Learn. Mem.* **2** 161–77
- [24] Rose J K, Kaun K R and Rankin C H 2002 A new group-training procedure for habituation demonstrates that presynaptic glutamate release contributes to long-term memory in *Caenorhabditis elegans* *Learn. Mem.* **9** 130–7
- [25] Ezzeddine Y and Glanzman D L 2003 Prolonged habituation of the gill-withdrawal reflex in *Aplysia* depends on protein synthesis, protein phosphatase activity, and postsynaptic glutamate receptors *J. Neurosci.* **23** 9585–94
- [26] Amano S, Kitamura K and Hosono R 1999 Hierarchy of habituation induced by mechanical stimuli in *Caenorhabditis elegans* *Zool. Sci.* **16** 423–9
- [27] Wang D and Ewert J P 1992 Configurational pattern discrimination responsible for dishabituation in common toads *Bufo bufo*: behavioral tests of the predictions of a neural model *J. Comp. Physiol.* **170** 317–24
- [28] Davis M 1970 Effects on interstimulus interval length and variability on startle-response habituation in the rat *J. Comp. Physiol. Psychol.* **72** 177–92
- [29] Staddon J E R 1993 On rate-sensitive behavior *Adapt. Behav.* **1** 421–36
- [30] Broster B and Rankin C H 1994 Effects of changing interstimulus interval during habituation in *Caenorhabditis elegans* *Behav. Neurosci.* **108** 1019–29
- [31] Giles A C and Rankin C H 2009 Behavioral and genetic characterization of habituation using *Caenorhabditis elegans* *Neurobiol. Learn. Mem.* **92** 139–46
- [32] Rankin C H and Broster B S 1992 Factors affecting habituation and recovery from habituation in the nematode *Caenorhabditis elegans* *Behav. Neurosci.* **106** 239–42
- [33] Rose J K and Rankin C H 2001 Analyses of habituation in *Caenorhabditis elegans* *Learn. Mem.* **8** 63–9
- [34] Rankin C H 2000 Context conditioning, in habituation in the nematode *Caenorhabditis elegans* *Behav. Neurosci.* **114** 496–505
- [35] Staddon J E R 2005 Interval timing: memory, not a clock *Trends Cogn. Sci.* **9** 312–4
- [36] McNamara A M, Magidson P D, Linster C, Wilson D A and Cleland T A 2008 Distinct neural mechanisms mediate olfactory memory formation at different timescales *Learn. Mem.* **15** 117–25
- [37] Alonso L, Moreno R, Vázquez M and Santacreu J 2005 Simulation of the filtering role of habituation to stimuli *Span. J. Psychol.* **8** 134–41



- [38] Steidl S, Rose J K and Rankin C H 2003 Stages of memory in the nematode *Caenorhabditis elegans* *Behav. Cogn. Neurosci. Rev.* **2** 3–14
- [39] Bekey G A 2005 *Autonomous Robots: From Biological Inspiration to Implementation and Control* (Cambridge, MA: MIT Press)
- [40] Floreano D and Mattiussi C 2008 *Bio-Inspired Artificial Intelligence: Theories, Methods and Technologies* (Cambridge, MA: MIT Press)
- [41] Belatreche A 2010 *Biologically Inspired Neural Networks: Models, Learning, and applications* (Berlin: VDM Verlag)
- [42] Gerstner W and Kistler W 2002 *Spiking Neuron Models: Single Neurons, Populations, Plasticity* (Cambridge: Cambridge University Press)
- [43] Izhikevich E M 2003 Simple model of spiking neurons *IEEE Trans. Neural Netw.* **14** 1569–72
- [44] Maass W 1997 Networks of spiking neurons: the third generation of neural network models *Neural Netw.* **10** 1659–71
- [45] Maass W and Bishop C M (eds) 1999 *Pulsed Neural Networks* (Cambridge, MA: MIT Press)
- [46] Di Paolo E A 2002 Spike timing dependent plasticity for evolved robots *Adapt. Behav.* **10** 243–63
- [47] Wang X, Hou Z-G, Zou A, Tan M and Cheng L 2008 A behavior controller based on spiking neural networks for mobile robots *Neurocomputing* **71** 655–66
- [48] Stanley J C 1976 Computer simulation of a model of habituation *Nature* **261** 146–8
- [49] Innis N K and Staddon J E R 1989 What should comparative psychology compare? *Int. J. Comp. Psychol.* **2** 145–56
- [50] Wang D L 1993 A neural model of synaptic plasticity underlying short-term and long-term habituation *Adapt. Behav.* **2** 111–29
- [51] Staddon J E R and Higa J J 1996 Multiple time scales in simple habituation *Psychol. Rev.* **103** 720–33
- [52] Staddon J E R 2001 *Adaptive Dynamics: The Theoretical Analysis of Behavior* (Cambridge MA: MIT Press)
- [53] Gutierrez-Osuna R and Gutierrez-Galvez A 2003 Habituation in the KIII olfactory model with chemical sensor arrays *IEEE Trans. Neural Netw.* **14** 1565–8
- [54] Marsland S 2009 Using habituation in machine learning *Neurobiol. Learn. Mem.* **92** 260–6
- [55] Marsland S, Nehmzow U and Shapiro J 1999 A model of habituation applied to mobile robots *Proc. TIMR 99 'Towards Intelligent Mobile Robots' (Bristol) Report UCMS-99-3-1*
- [56] Stiles B W and Ghosh J 1997 Habituation based neural networks for spatio-temporal classification *Neurocomputing* **15** 273–307
- [57] Lara R and Arbib M A 1985 A model of the neural mechanisms responsible for pattern recognition and stimulus specific habituation in toads *Biol. Cybern.* **51** 223–37
- [58] Broekens J, Koesters W A and Verbeek F J 2007 Affect, anticipation, and adaptation: affect-controlled selection of anticipatory simulation in artificial adaptive agents *Adapt. Behav.* **15** 397–422
- [59] Marsland S, Nehmzow U and Shapiro J 2005 On-line novelty detection for autonomous mobile robots *Robot. Autonomous Syst.* **51** 191–206
- [60] Chang C 2005 Using sensor habituation in mobile robots to reduce oscillatory movements in narrow corridors *IEEE Trans. Neural Netw.* **16** 1582–9
- [61] Sirois S 2005 Hebbian motor control in a robot-embedded model of habituation *Proc. IEEE Int. Joint Conf. on Neural Networks* **5** 2772
- [62] Chiel H J and Beer R D 1997 The brain has a body: adaptive behavior emerges from interactions of nervous system, body and environment *Trends Neurosci.* **20** 553–7
- [63] Chrisley R 2003 Embodied artificial intelligence *Artif. Intell.* **149** 131–50
- [64] Pfeifer R 2007 *How the Body Shapes the Way We Think: A New View of Intelligence* (Cambridge, MA: MIT Press)
- [65] Ziemke T 2003 What's that thing called embodiment? *Proc. 25th Annual Meeting of the Cognitive Sci. Society (Boston, MA)* ed R Alterman and D Kirsh (Mahwah, NJ: Lawrence Erlbaum) pp 1134–9
- [66] Mataric M 2007 *The Robotics Primer* (Cambridge, MA: MIT Press)
- [67] Wang P and Goertzel B (eds) 2007 Introduction: aspects of artificial general intelligence *Advances in Artificial General Intelligence: Concepts, Architectures and Algorithms* ed B Goertzel and P Wang (Amsterdam: IOS Press)
- [68] Johnston J 2008 *The Allure of Machinic Life: Cybernetics, Artificial Life, and the New AI* (Cambridge, MA: MIT Press)
- [69] Sendhoff B, Körner B, Sporns O, Ritter H and Doya K 2009 *Creating Brain-like Intelligence: From Basic Principles to Complex Intelligent Systems* 1st edn (Berlin: Springer)
- [70] Krohs U and Kroes P 2009 *Functions in Biological and Artificial Worlds: Comparative Philosophical Perspectives* The Vienna Series in Theoretical Biology (Cambridge: MIT Press)
- [71] Cyr A, Boukadoum M and Poirier P 2009 AI-SIMCOG : a simulator for spiking neurons and multiple animat's behaviours *Neural Comput. Appl.* **18** 431–46
- [72] Cyr A, Thériault F and Boukadoum M 2012 NeuroSim: a virtual 3D-world to investigate the intelligence phenomenon within the perspective of bio-inspired robotic agents *Virtual Worlds - Artificial Ecosystem and Digital Art Exploration* ed S Bornhofen, J-C Heudin, A Lioret and J C Torrel (Paris: Science eBook) pp 167–85
- [73] Abbott L F and Kandel E R 2012 A computational approach enhances learning in aplysia *Nature Neurosci.* **5** 178–9
- [74] Cyr A and Boukadoum M 2012 Classical conditioning in different temporal constraints: an STDP learning rule for robots controlled by spiking neural networks *Adapt. Behav.* **20** 257–72

## ANNEXE C

### ARTICLE 2 : CONDITIONNEMENT CLASSIQUE

#### ADAPTIVE BEHAVIOR

# Classical conditioning in different temporal constraints: an STDP learning rule for robots controlled by spiking neural networks

André Cyr and Mounir Boukadoum

## Abstract

This work investigates adaptive behaviours for an intelligent robotic agent when subjected to temporal stimuli consisting of associations of contextual cues and simple reflexes. This is made possible thanks to a novel learning rule based on spike-timing-dependent plasticity and embedded in an artificial spiking neural network serving as a brain-like controller. The subsequent bio-inspired cognitive system carries out different classical conditioning tasks in a controlled virtual 3D-world while the timing and frequency of unconditioned and conditioned parameters are varied. The results of this simulated robotic environment are analysed at different stages from stimuli capture to neural spike generation and show extended behavioural capabilities by the robot in the temporal domain.

## Keyword

Spike-timing-dependent plasticity, spiking neuron, bio-inspired robotics, classical conditioning

## 1 Introduction

One consensual feature of intelligent behaviour is adaptation through experience. Because of this, learning and memory have received wide attention in the neuroscience literature and certain neural structures appear to play an essential role in these processes when seen from a high-level perspective (Rudy, 2008). At a smaller biological scale, the synapse appears to be of paramount importance, underlying physical changes that drive these learning and memory processes. In particular, the synaptic information exchange and modulation effects are thought to be the underlying mechanisms of the plasticity phenomenon (Cowan, Südhof & Stevens, 2001). The biochemical synapse can thus be viewed as a highly adapted dynamical device that can distinguish between local activities resulting from distinct temporal events occurring at both ends of a neuron. These features allow the synapse to act as a plastic gate exhibiting coincidental detection and directional synaptic change in response to a temporal order of pre- and post-spikes (Markram, Lübke, Frotscher & Sakmann, 1997), where the precise spike timing and frequencies are among the elements causing these physical changes. For instance, it is known that repetitive spike pairing induces short- and long-term potentiation or depression


(LTP/LTD) of the synapse, hence showing experimental correlates of the Hebbian theory. These processes are thought to lead to the elementary memory trace in the brain, formalized by synaptic weights (Letzkus, Kampa, & Stuart, 2007).

The temporal correlations of pre- and post-synaptic spiking activity are known to result partially in a spike-timing-dependent plasticity (STDP) process (Bi & Poo, 1998). However, there remain pending issues with the theoretical STDP concept (Pfister, 2006) and its future directions (Markram, Gerstner, & Sjöström, 2011). One is its non-linear regulation (Shouval, Wang, & Wittenberg, 2010) and saturation mechanisms to prevent the endless build-up of weights from repetitive associations. Natural burst spiking and individual spike frequencies also influence the STDP learning curves. The introduction of homeostatic boundaries and down-scaling processes from other parameters may also restrain weight growth. A biologically plausible

Computer Science Department, Université du Québec à Montréal, Montréal, Québec, Canada

## Corresponding author:

André Cyr, Computer Science Department, Université du Québec à Montréal, Montréal, Québec, Canada  
Email: andre.cyr1@videotron.ca

Adaptive Behavior  
20(4) 257–272  
© The Author(s) 2012  
Reprints and permissions:  
sagepub.co.uk/journalsPermissions.nav  
DOI: 10.1177/1059712312442231  
adb.sagepub.com  




mechanism has been proposed (Roberts & Leen, 2010) to explain the down-regulation synaptic weight cause from non-correlational spike time-pairing. As a result, the current computational synaptic models appear to display a proper balance between plasticity levels in response to continuous dynamical input-output information, and sufficient stability for the memory recall process.

Recently, the STDP features have been re-examined in the light of the large experimental data accumulated on the subject (Caporale & Dan, 2008). The ensuing models point to various secondary messenger pathways as causes for the dynamics of STDP (Bi & Poo, 2001; Bohte & Mozer, 2007; Dan & Poo, 2006; Kistler & van Hemmen, 2000; Roberts & Bell, 2002; Song, Miller, & Abbott, 2000), that mostly impact the synaptic weights by dynamically modulating the width and amplitude of the critical time window from the coincidental pre-post or post-pre spike response. These findings call for new integrative models of synaptic effects that account for both the temporal and rate constraints of spike coding. The artificial spiking neural network (ASNN) (Christodoulou, Bugmann, & Clarkson, 2002; Gerstner & Kistler, 2002; Izhikevich, 2003; Maass & Bishop, 1999) is one such unitary spike resolution model (Bohte, 2004) that has been popular for cognitive and robotic modelling in recent years. Synaptic plasticity (Mehrtash et al., 2003) and more specifically STDP learning have also been combined with ASNN to investigate biological networks (Bush, Philippides, Husbands, & O'Shea, 2010; Florian, 2010; Kempter, Gerstner, & Hemmen, 1999; Soula, Alwan, & Beslon, 2005; Wang, Hou, Zou, Tan, & Cheng, 2008). However, efforts to combine STDP, LTP/LTD and ASNNs for robotic applications are less common (Alnajjar & Murase, 2008; Alnajjar, Zin, & Murase, 2009; Arena, De Fiore, Patane, Pollino & Ventura, 2009; Di Paolo, 2002). A possible explanation for this is that the added layer of temporal constraints from the sensory stimulus protocol (Froemke & Dan, 2002) calls for integrative models accounting for the whole chain of delays that exist between stimuli, from their detection by transducers, to the propagation of internal spikes patterns in the neural circuit, to the final motor behaviour.

This paper contributes to the subject by examining associative learning, a well-known aspect of classical conditioning (Schmajuk, 2010) that has already been studied with ASSN (Lieu & Shapiro, 2007), when reduced to the level of a small neural circuit that operates under a STDP rule. The objective is to expand the temporal domain of adaptive behaviour in the robotic field by exploring the learning responses from different associative temporal stimuli contexts. A bio-inspired ASSN model with a novel STDP learning function is used for this purpose. The proposed brain-like controller is embedded in a virtual robot that evolves in a

dynamical 3D-world to address the following question: how does a robot learn associations between external temporal stimuli patterns and modulate its reflex behavioural repertoire as a result? As a partial answer, we propose a STDP rule that modulates synaptic strength from the occurrences of pre-post or post-pre spikes, temporally correlated within a critical time window, and also integrates the influence of traditional paired-spikes timing with paired and individual spikes frequencies within the STDP rule, thus leading to a temporal all-to-all spikes interaction model.

Current STDP models can explain experimental data under controlled protocols, but remain incapable to do so for random spiking patterns that arise in natural situations (Dan & Poo, 2006). These models fail also to encompass the entire diversity of synaptic plasticity and related experimental data for all known STDP attributes. A step to address these shortcomings is the development of a STDP rule where each individual pre- or post-synaptic spikes pattern, or timed combination of paired pre-post or post-pre spikes, triggers a unique time-delayed event that modifies the synapse. Depending on the paired-spikes timing, spikes pairing frequencies and individual pre- and post-spikes frequencies (Verhoog & Mansvelder, 2011), different courses of action can take place to produce synaptic changes in reaction to temporally correlated neural inputs and outputs. The post-spike effect has been investigated in experimental studies, but few reports exist in literature on a pre-spike effect (Shouval, Wang & Wittenberg, 2010). In this work, we propose to use paired-spikes timing, spikes pairing and individual pre- and post-spike frequencies as features in an all-to-all STDP model, where the synaptic changes integration influences the artificial agent's learning curve at the behavioural level. When spike frequencies are put in correlation with external stimuli patterns occurring over different timescales, they may help the agent categorize or respond differently to certain inputs, as in classical conditioning observations. Communication and context cues in navigation are examples of situations where temporal structured information could potentially be meaningful to AI agents.

In the simulations we report, the AI agent reflects two poles: relevant biological facts at the temporal scale and actual robotic possibilities. This enables comparisons between AI and natural agents without seeking to achieve the strict accuracy of the latter. In this sense, biology becomes a tool for building bio-inspired AI agents (Floreano & Mattiussi, 2008), and there may be also advantages to the reverse scenario.

The balance of this paper is organized as follows: first we extend a standard STDP rule by merging non-linear effects, saturation mechanisms, relative spike timing, spike pairing frequency, and individual pre- and post-spike frequency variables. Second, we include the novel STDP rule in an ASNN model. Third, the



ASSN acts as a brain-controller for a simulated robot equipped with sensors and effectors, enabling it to perceive and act in a virtual 3D-world. Fourth, different temporal sets of stimulus spikes are applied into the virtual world while the robot is at rest and when it moves freely. We present the results of the learning function at the behavioural level of a minimal cognitive system. Finally, we challenge the simulated robot in traditional Pavlovian tasks.

## 2 Methodology

The AI agent is a simulated robot (replication of Khepera III, K-Team co.) using an ASSN as brain-controller. The robot is made to perform conditioning tasks within a virtual 3D world. Different stimuli protocols are applied to cover the different temporal aspects of the STDP learning rule. The framework is provided by AI-SIMCOG (Cyr, Boukadoum, & Poirier, 2009), a neural simulator software dedicated to investigating the global intelligence phenomenon. AI-SIMCOG can simultaneously handle multiple abstraction levels, from the detailed synaptic model to the behaviour observed in several autonomous mobile robots. The platform is largely inspired from low-level biological knowledge, neural information processing and the embodiment paradigm (Chiel & Beer, 1997), all applied to cognitive robotic and artificial life models, be they virtual or physical AI agents. An important feature of AI-SIMCOG is its ASSN editor, which includes built-in basic learning functions and the facility to translate the ASSN description into code for a physical robot.

To better understand the enhanced STDP rule, a brief survey of AI-SIMCOG's ASSN implementation is given in Section 2.1. Then, a description of the external shape and the neural architecture of the simulated robot follow in Section 2.2. Finally, the experimental protocol including the context and tasks accomplished by the virtual robot to demonstrate the strength of the STDP algorithm and its impact on the robot's behaviour are described in Section 2.3.

### 2.1 ASSN: internal dynamic description

The ASSN algorithm of AI-SIMCOG relies on repeated discrete-time cycles that use pre-stored sampled function shapes. The different variables are updated through pre-determined numerical look-up tables instead of using differential equations as done in standard leaky integrate-and-fire (LIF) ASSNs. The table contents are percentage scalar numbers derived from response curves and relevant constants representing the major phenomenological aspects of current neural models. An advantage of this equivalent method is a significant gain in computational efficiency, since no complex calculations are needed online, and the

ability to use arbitrary response shapes without the need for prior analytic formulas; the disadvantage is a potential loss of accuracy in comparison with using continuous time functions, and the lack of closed form solutions. Depending on the size and complexity of the neural architecture under investigation, the duration of one step cycle can be tuned to one-millisecond resolution for real-world applications, while taking into account the limitations of the time buffers of physical sensors and the minimal CPU efficiency.

The main property of AI-SIMCOG's neurons is based on the membrane potential's dynamic behaviour, varying according to various stored non-linear curves in response to the integration of multiple inputs. An input consists of simulated excitatory or inhibitory currents inducing a local post-synaptic potential variation. The artificial electrical current comes from different sources: external stimulus processes perceived from transducers attached to sensory-neurons, neuromodulators, neural spikes and virtual electrodes. The thresholds are fixed by the user and, when crossed, lead to an action potential of a predetermined shape. The time-frame for the refractory period variable is settable. The after-spike return value is another parameter, creating a momentary hyperpolarized state and modulating the behaviour of the membrane potential curve. An internal neural leak feature is also included, allowing the membrane potential to relax to a steady homeostatic resting state if desired, as done in ASSN models with LIF neurons. A stronger leak for neurons is an option, introducing continuous intrinsic spike oscillations or endogenous pacemaker beats. This characteristic adds a useful dynamic feature to the ASSN in situations lacking sensory input, as often seen at the start of a simulation. The initial conditions for all ASSN constants and variables are manually set,<sup>1</sup> but could also be tuned by computational algorithms or evolutionary strategies.

Neurons connect via single or multiple synaptic links for arbitrary network topologies. The synaptic properties determine the effect on the post-synaptic unit: a positive or negative inward current. AI-SIMCOG's synaptic models allow simultaneous integration of multiple learning rules, occurring over different temporal scales. The static number found in traditional ASSN models for the weight value is replaced by a row in the built-in look-up tables, operating on a determined time interval. Therefore, a specific synaptic weight corresponds to a percentage number that multiplies a vector in the table, and its instantaneous value is one of the resulting vector elements. When a weight vector is solicited, it mimics a standard alpha curve (modelled as the difference of two exponential shapes) for an excitatory or inhibitory post-synaptic potential, EPSP or IPSP, while one of the weight vectors in the look-up table (synaptic matrix) will provide the post-synaptic potential variation sequence, depending on the



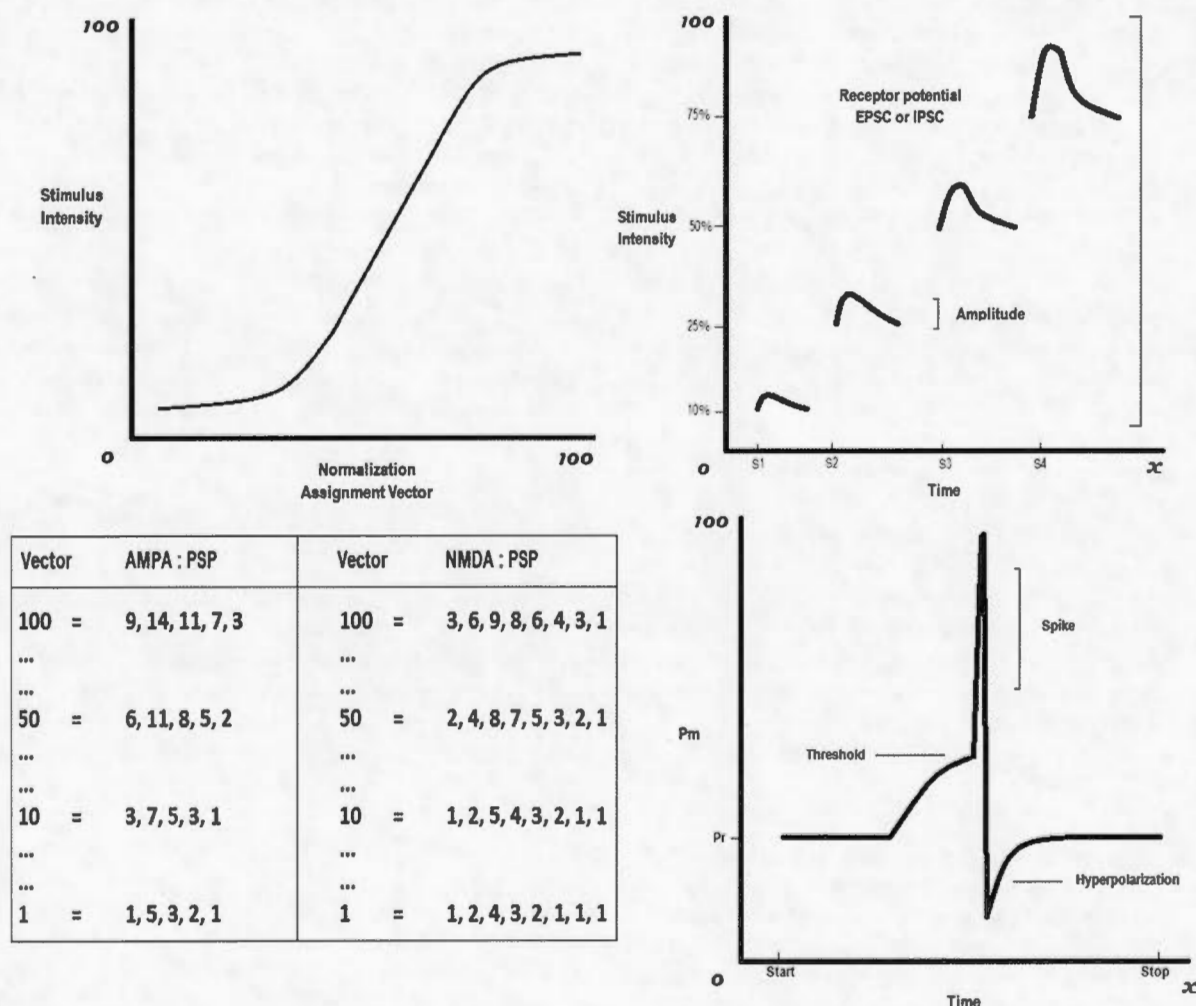
**Figure 1.** AI-SIMCOG general description.

intensity range of the local membrane potential variation produced by synaptic activity. These are simulated by local current changes, summed up and integrated with the membrane potential, eventually crossing a determined threshold for the emission of spikes. When the output spikes target synapses, the effect is to generate a fluctuation in the membrane potential of the receiving post-synaptic neuron, depending on the synaptic weight and the attached learning rules. The

neural network responses are further processed as inputs to various targets (interneuron, motoneuron and effector), eventually producing actions in the real or virtual world. Figure 1 summarizes the AI-SIMCOG general action sequence and a given internal configuration is illustrated in Figure 2.

## 2.2 STDP model

The main parameters involved in this proposed STDP model are the spike timing, the frequency of paired spikes, the individual frequency of pre- and post-spikes and dynamic sets of synaptic weight vectors associated with these variables. A voltage variation scheme is used for the determination of the spike timing as well as the



**Figure 2.** Artificial spiking neural networks (ASNNs) primary internal mechanisms. Top left: Capture of a discrete external stimulus by a transducer component and normalization after sigmoid processing, resulting in a single scalar percentage number. Top right: Percentage stimulus intensity corresponding to a pattern of receptor potential variation of an alpha function shape (S1–S4) integrated in the membrane potential of the sensory neuron. Bottom left: Normalized synaptic weights (1–100), where each weight corresponds to vectors composed of scalar numbers varying in time and amplitude. These vectors are the post-synaptic potential (PSP) to integrate and depend on the current state of the membrane potential. Bottom right: Typical membrane potential (Pm) curve starting at the rest potential (Pr), non-linearly ascending to reach the threshold and provoking a spike of a given absolute refractory period. It is followed by a short hyperpolarization state, and a gradual return to the rest potential.



representation of the synaptic weights. Therefore, this STDP rule could be qualified as a membrane potential-dependent process. All kinetics functions used in the algorithm are based on the qualitative shape of the functions found in experimental data.

The STDP algorithm is based on multiple conditions (Figure 3, left) pointing to a matrix of numerical vectors and applied online at each time step. Vectors are built with percentage scalar numbers representing the goal of the expected synaptic change. The first rule implies a synaptic weight variation in response to the temporal offset between pre- and post-synaptic firing (Figure 3, right). Here, the pre-synaptic neuron is designated A and the post-synaptic neuron B. If neuron A fires before neuron B (causal mode) in a delimited narrow time window, the synaptic weight increases with some variation from the current weight depending of the exact timing. The closer the paired spikes are in time, the more the synaptic strength will change. This non-linear temporal relation constitutes the first index of the synaptic matrix embedded in the STDP learning rule. In the opposite situation, or acausal mode, if neuron B fires before neuron A, the synaptic weight decreases. The resultant curve shape is not the perfect mirror image of the causal mode and a slight asymmetry exists between the two responses. The right side of Figure 3 highlights these relations between synaptic changes and spikes timing from a STDP paradigm (based on a schematic drawing of experimental data from Bi & Poo, 1998).

Once a spike is detected, an internal clock starts for the neuron emitter (A) with a reset delay timeframe of 100 cycles. If a target neuron (B) emits a spike within the relevant time period (causal mode), then the applied STDP rule shifts momentarily the synaptic vector pointer toward a higher number which increase the effect of eventual EPSPs or IPSPs. On the contrary, if the pre-synaptic neuron (A) receives a spike from the post-synaptic neuron (B) prior to its spike emission, meaning a non-correlational association, then the synaptic vector pointer or weight will decrease. Nothing

happens when a spike occurs out of the time boundaries (which can be set to correspond to any desired times), exceeding the expected biological synaptic behaviour. Thus, for each pair of spikes occurring on the defined temporal window, regardless of the causal or acausal mode, a unique scalar number is dynamically attributed from the precise timing of the spikes. This number is converted to one of five predetermined categories, corresponding to the first index of the synaptic vector or the desire weight modulation to come.

To reach any reassigned synaptic weight in the matrix, higher or lower, the response is not instantaneous but occurs within a predefined period of cycles (positive exponential and decay functions set arbitrary on five cycles). This latency parameter relies on biological plausibility where synaptic changes occur on varying timescales and different temporal lengths, from milliseconds to minutes, depending on the underlying biochemical mechanism route (may reflects an early phase of LTP/LTD). Also, if successive correlations of different timings of paired spikes occur, then the integration of this part of the STDP rule is done by averaging all processes online.

The second index refers to the time interval between each pair of spikes or the effect of different frequencies of paired spikes on the synaptic weight variation response. Thus, the different paired spikes contribute unequally to the dynamical synaptic process, depending on a temporal order parameter or the nearest-neighbour interaction. For example, two pairs of spikes with the same timing between the pre- and post-spike, occurring with 20 cycles of difference will produce more synaptic changes than when they are separated by 50 cycles. This frequency-STDP feature stands on the same logic for traces (voltage) left by a specific event. The five temporal categories are the same as the timing parameter.

In addition to the timing of paired spikes and frequency of paired spikes, we also model synaptic plasticity dependence for individual spike frequencies arising independently on both synaptic sides. The pre-synaptic spike frequencies may alter the synapse as found in

**If paired spikes occurred within a define time steps**

**Set Index1 = Paired-Spike Interval**

**Set Index2 = Mean of frequencies (Inter-Paired-Spike Interval)**

**Set Index3 = Mean of individual pre and post-spike frequencies**

**Categorize each index = [1-5]-[6-10]-[11-25]-[26-50]-[51-100]**

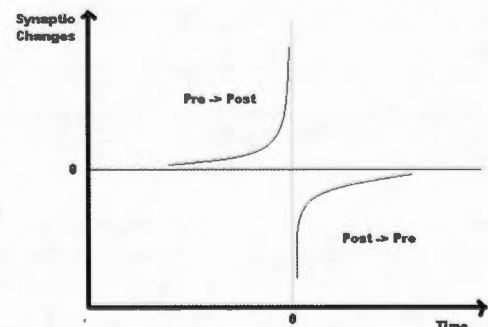
**Depending on causal or acausal mode**

**-> Point to a specific vector number in the matrix (synaptic change)**

**Adjust the current synaptic weight with the synaptic target**

**If there is a forgetting term (recovery function)**

**-> Adjust the current synaptic weight to the original value**



**Figure 3.** Left: summary of the spike-timing-dependent plasticity (STDP) algorithm. Right: general qualitative curves representing the synaptic changes as a function of timed pre-post and post-pre spikes, leading to long-term potentiation or depression (LTP or LTD) for specific protocols. Similar curves are obtained for paired-spikes interval and the individual spike frequency parameters.

short-term plasticity concept, whether it is with paired-pulsed facilitation, post-tetanic potentiation, and sensitization or habituation mechanisms. The idea is that several consecutive pre-synaptic spikes influence the synapse, even when they are not paired with a post-spike. The post-synaptic spike frequency also modulates the synapse from back-propagation current and calcium concentration modification. For example, if neuron A emits two consecutive pre-spikes followed by the emission of a single post-spike from neuron B within the critical time window, then the synaptic weight will increase more than in the case of a unique pair of spikes. This situation is designated by a 2:1 or pre-pre-post spike condition. Consecutive emission of pre- or post-spikes will be saturated by their own refractory period. The highest firing frequency results in the strongest synaptic change. For example, if neuron A produced two consecutive spikes within 100 cycles, in one case separated by eight cycles and in the other by 80 cycles, the first condition will have more impact on the synaptic change if it is followed by any spike of neuron B occurring within the timeframe limit of the STDP function. The same logic applies if neuron B fires twice after the previous emission of a spike from neuron A.

Thus, if more pre- or post-spikes occur, more synaptic effect is expected depending on the timing and individual spike frequency. This factor constitutes the third index of the synaptic matrix, varying between 0 and 100 and segmented into the same five percentage number categories as in the paired spikes timing. This additional temporal relation feature also sustains asymmetric curves between causal and acausal modes. Spike frequencies can change instantly during a simulation and a summation is averaging online, integrating all previous spike frequencies occurring on both neural side elements. This individual spike frequency variable gives now a coherent interpretation to manage any neural spike pattern, including possible future add-ons in the ASNN in the form of neural burst features or noise functions.

After processing these three indices of relative equivalent impact on the synaptic variation, the STDP rule uses a recovery function to backtrack to the original weight.<sup>2</sup> Thus, a non-linear decay starts whenever a weight change sequence is launched, with an arbitrary time constant of 200 cycles. Depending on the target of this modulation, the backtrack time length takes two times longer than the raise of the previous changes (the period is arbitrary set). Finally, this STDP model could be applied to other computational paradigms, but is better understood as an integrative part of the ASNN algorithm closely linked to an AI agent.

### 2.3 Robot external structure, internal architecture and experimental protocol

The virtual robot (Figure 4) had a circular body 30 cm in diameter and was equipped with several types of

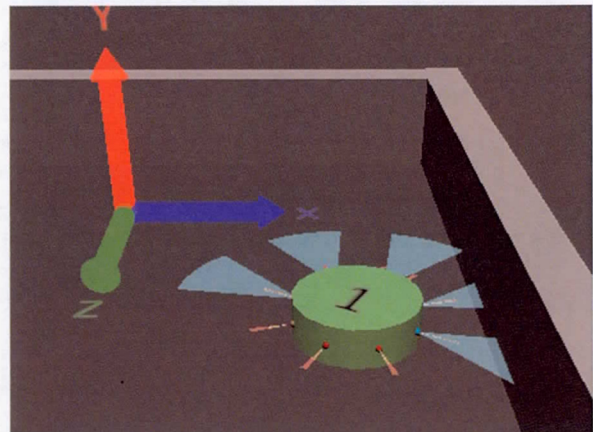


Figure 4. Robot external structure.

sensors, represented as virtual transducers attached to sensory neurons in the ASNN editor. To interact with its environment, the robot starts with pre-wired reflexes, allowing basic locomotion patterns and active perceptual abilities. To allow perception, five long-range ultrasound sensors are implemented around the robot at 0, 45, 90, 270 and 315°. The field range of these transducers is tuned to between 20 and 30 cm, with a maximal response when the robot is close to an object. The robot is also equipped with nine short-range infrared sensors calibrated to trigger a response between 0 and 10 cm, with a maximal response when the robot is close to an obstacle. These sensors are also capable of reading the ambient light in passive mode, and thus provide nine additional sensors in the visual modality. One of these ambient light sensors is founded under the robot, giving the possibility to detect difference in the colour pattern of the floor. In the next simulations described below, the ambient light sensors serve to perceive the neutral conditioning stimulus (CS), presented as different temporal light patterns. The visual sensor modality does not trigger a particular behaviour before the conditioning tasks.

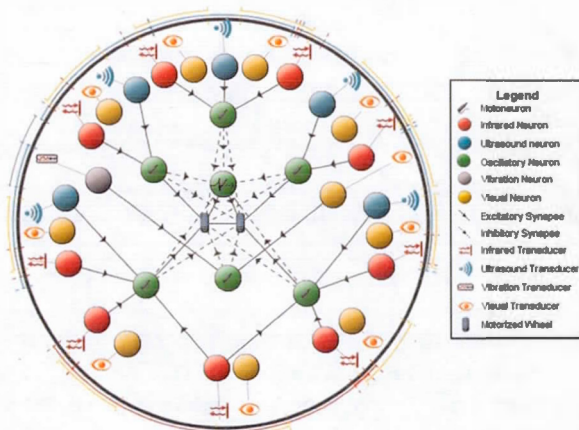
The robot is also equipped with a vibration sensor allowing perception of the stimuli from the world. This modality serves as unconditioned stimulus (US) in the scenario. As such, any vibration triggers an innate aversive behaviour in the form of a backward movement, the unconditioned response (UR). The learning task consists of also associating the CS with this fear reflex (negative tropism). The goal is to have a conditioned response (CR) to a neutral stimulus develop, by associating the repetitive occurrence of a CS followed by a US. This unsupervised learning process implies a forward conditioning with various temporal delays between two homogeneous simple external stimuli. Consequently, the CS precedes the US, both stimuli leaving an observable trace or memory where the response of the association is a synaptic change. This is



a typical form of classical conditioning. In an AI learning context, we conveniently assume a temporal shift from the psychological level (Pavlovian conditioning) to the molecular level (STDP), considering that the time delay between the CS and the US is adjusted to be consistent with the STDP interval window (These temporal parameters could be tuned differently). In summary, an external set of stimuli generates an internal neural spike pattern and through the STDP rule, modulates the behaviour of the robot from past experience. The sequence is predictable because, as synaptic changes cumulate or reinforce the updated weight, it turns eventually the neutral response of the CS into a CR: an observable backward movement when a light stimulus is perceived alone.

The ASNN architecture (Figure 5) uses one sensory neuron per transducer, connected to a motoneuron. Motoneurons are connected to effectors, the motorized wheels. The ASNN adds an intrinsic leaky neuron that periodically spikes and leads to continuous forward locomotion behaviour. For the sake of simplicity, only the synapse between the visual sensory neuron and the motoneuron is activated for plasticity. The ASNN architecture and the starting synaptic weight matrix were intuitively hand-coded. At this stage, there was no point achieving some efficient or fitness criteria from evolutionary strategies like genetic algorithms to fix the parameters, though this could have been done externally. In the same vein, randomizing inputs such as external physical stimuli, spike trains generated by the neurons or the starting synaptic weight matrix were not pursued as they frequently caused a non-interpretable chaotic behaviour of the robot.

The virtual world consists of a 3D arena of 6×6 m delimited by walls, in which the robot is exposed sequentially to the CS follow by the US. In the first part of the experimental protocol, external stimuli patterns are achieved as discrete uniform flashes of light

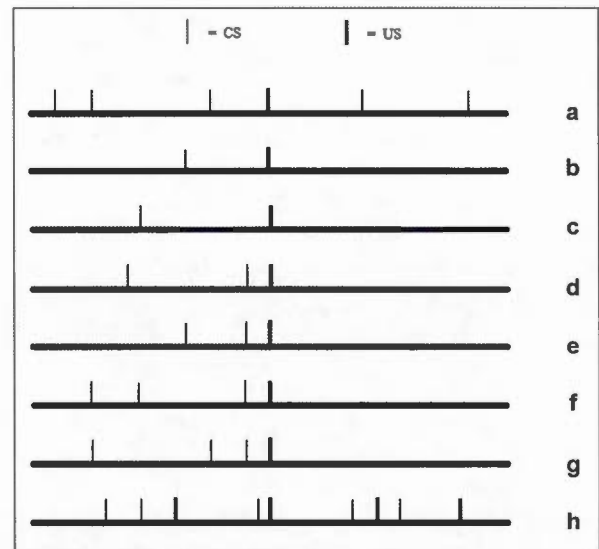


**Figure 5.** Robot internal neural architecture with some synaptic links, based on Khepera III.

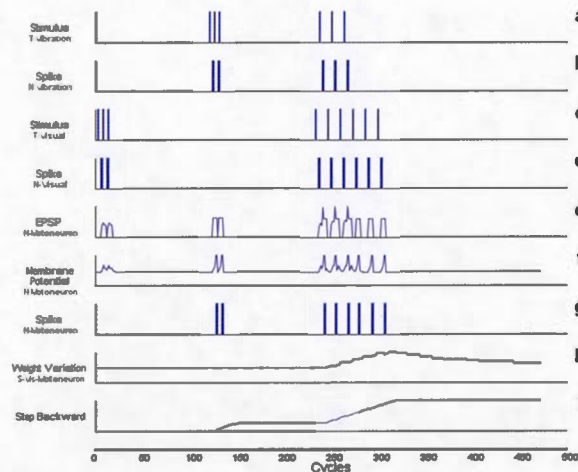
and vibration emissions delivered into the entire world with the robot at rest. Varying frequency patterns are set for the CS and US tested under both causal and acausal conditions (Figure 6), in order to demonstrate the STDP features. In the second part of the protocol, the illustration of a specific conditioning task supported by the STDP mechanism during normal locomotion behaviour is provided. The experiment section ends with a web link to a real-world conditioning application using a modified Khepera III robot, which reproduces the behaviour observed in the virtual world.

### 3 Simulations and results

The first section to follow concerns a robot at rest, with different combinations of CS and US applied along a trace forward conditioning, with one stimulus producing one spike. The various temporal patterns show the main features of the STDP rule, from the transducer level to the robot's behavioural response. Because of

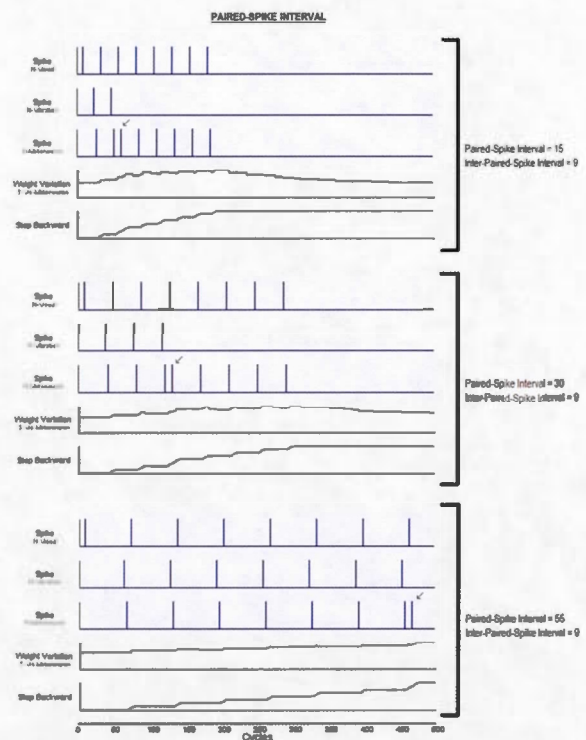


**Figure 6.** Examples of external stimulus patterns generating spiking activity. (a) Few conditioning stimuli (CS) occurring before and after a single unconditioned stimulus (US). (b) One CS applied before a US, expecting a positive association. (c) This CS is likely to evoke a weaker association, because of an increase time length, compared with the trial in (b). (d) Two consecutive CS occur before a US, reinforcing the association compared with (b). (e) Stronger association compared with (d), because of the higher individual frequency and the short timing of the CS. (f) Higher frequency of the CS leads to a stronger association than in previous traces. (g) The same frequency of the CS and the same time frame, but with a shorter timing for the middle stimulus, resulting in a stronger association compared with (f). (h) Complex temporal pattern of CS and US with hard to predict association, but still computationally tractable from the presented spike-timing-dependent plasticity (STDP) rule and its related temporal features.



**Figure 7.** After testing the reflex response from the unconditioned stimulus (US; vibration) and the neutrality of the conditioning stimulus (CS; visual) in the outer time limit of the spike-timing-dependent plasticity (STDP) window, a short sequence of CS followed from few cycles in time by US is launched at cycle 250, allowing the CS alone to trigger a backward movement after a few pairs of spikes. The external stimuli are caught by transducers (T) producing a correlated spike pattern (slightly time delayed) from their sensory neurons (N) (a–d). These induced spikes generate excitatory post-synaptic potentials (EPSP) (e), which are integrated into the membrane potential of the motoneuron (f). Spikes are produced from the associative motoneuron (g). The cumulative positive process is observed from the weight variation when paired spikes are recorded at the synapse (S) between the visual neuron and the motoneuron (h). Whenever the motoneuron launches spikes, it induces a backward movement (i).

the pre-determined wiring and synaptic weights in the ASNN, the robot exhibits a reflex backward movement whenever it perceives a vibration (US) (Figure 7). This action represents the necessary trace within the embodiment paradigm, which served to monitor all the behavioural responses from the AI agents. As time progresses, no movement is observed when the robot is exposed to several light stimuli (CS) without encountering a vibration stimulus (US) within the critical time window. When the CS and the US occur during the same absolute range of 100 cycles, an associative temporal relation is made at the corresponding synapse according to the STDP rule. Because the CS immediately precedes the US, the causal mode is triggered in the algorithm. These paired spikes lead to a positive synaptic change indicated by the STDP trace or weight variation, with modulation strength set by the spikes timing. When cumulated at the visual neuron, the repetitive synaptic changes eventually reach the level to trigger a backward movement (CR) from the perception of a single light stimulus, engaged through the associative motoneuron pathway. When the CS ceases, the synaptic weight returns to the original baseline because of a

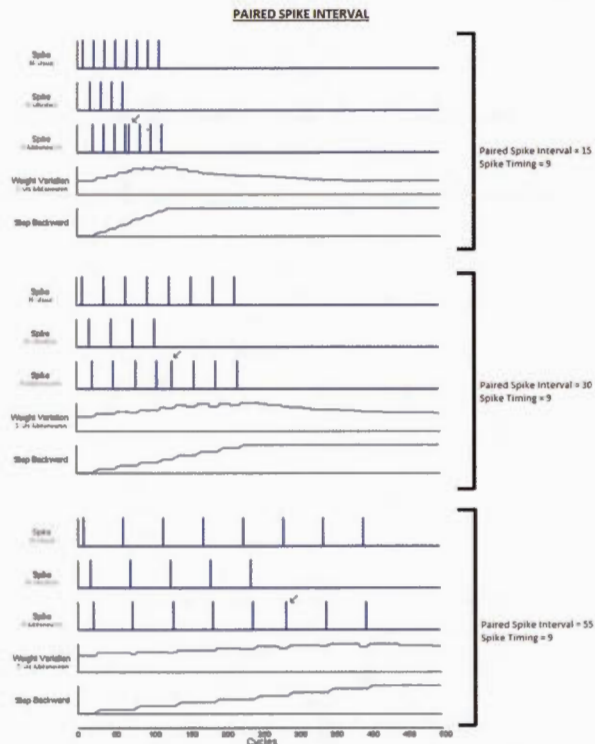


**Figure 8.** Three simulations with different paired-spikes intervals and a fixed inter-paired-spikes interval. In each case, a conditioning stimulus (CS; light flash) is paired with an unconditioned stimulus (US; vibration wave) to generate a single spike within the same temporal order. The associative motoneuron responds automatically (strong synapse) to each pre-synaptic spike incoming from the vibration sensory neuron. A weak synaptic weight between the motoneuron and the visual sensory neuron is set at the start of simulation, but repetitive causal association of the pre- and post-spike increase the weight until the visual sensory neuron generates spikes to the motoneuron. The arrow points to the new conditioned response (CR). The results confirm experimental data that the strongest synaptic weight variation comes from the smaller timing between the pre- and post-spike.

decay function. At the end of this simulation, further visual stimuli were unable to induce reflex backward movements (not shown in the graphic).

Figure 8 illustrates the effect of spike timing on synaptic plasticity and robotic behaviour. Three separate simulations were conducted with different values for the time difference between the pre- and post-spike (paired spike interval) with a fixed time delay between each pair of spikes (inter-paired-spikes interval). As expected, the results show that synaptic weight grows more rapidly when the paired spike timing is short. The CR is the result of consecutive increases of the synaptic weight at the visual neuron, inducing a spike response at the motoneuron. As a correlated result, the robot moves backward more rapidly in the smallest spike timing condition. Thus, starting from the same starting

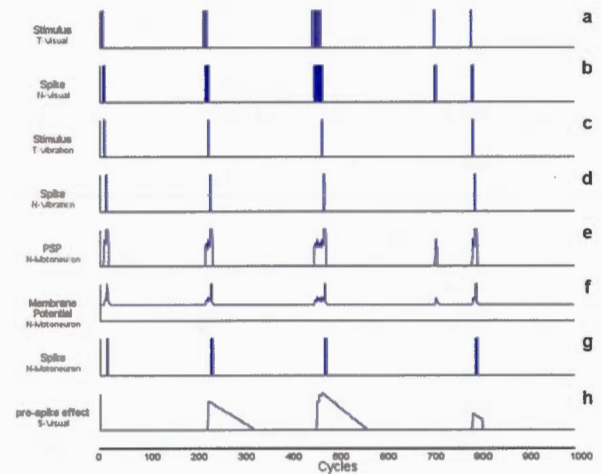




**Figure 9.** Three simulations highlight the effect of the time interval between each paired-spike on the synaptic changes and the robot's behaviour. All three simulations show that increasing the length interval induces a positive slope to reach the critical weight for the synapse between the associative motoneuron and the visual sensory neuron. The repetitive association of the external light stimulus (CS) preceding the external vibration stimulus (US) leads to a correlated temporal spike sequence at the neural component, resulting in a positive weight variation because of the spike-timing-dependent plasticity (STDP) rule and the particular feature of inter-paired-spikes interval management. The strongest variation occurred when the interval is close in time.

synaptic weight, it takes fewer associations to reach the critical threshold for spikes induction.

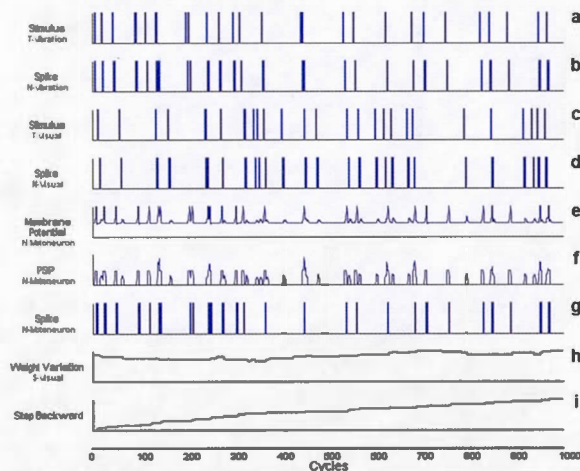
Figure 9 shows the effect of introducing different time intervals between repetitions of paired-spikes (inter-paired-spikes interval) with a constant paired-spikes interval of nine cycles between the pre- and post-neurons. The graphics shows the impact of different temporal sets of homogeneous time-paired spikes on the synaptic weight dynamics, favouring the smaller interval for greater synaptic modulation. When repetitive sequences of CS preceding a US are perceived, the association catches on from the STDP rule applied to the synapse between the motoneuron and the visual neuron, moving the weight toward a higher value. Again, when the critical point is reached, the CS alone elicits the desired learned fear response. Thus, higher frequencies of homogeneous paired spikes induced more synaptic change than a lower frequency.



**Figure 10.** Effect of different pre-spike frequencies on the parameter that affect the synaptic weight variation. The same logic applies to the post-spike frequency variable and the acausal mode. First, we show one pre-post spike association that does not trigger a response from the pre-spike effect parameter. It is followed by a 2:1 (pre-pre-post) trial around cycle 200, which induces a variation of the pre-spike effect variable. This variation increases non-linearly in the 4:1 situation. At the end, we offer another 2:1 trial, but where the timing of the first pre-spike is more distant than in the previous condition, with less effect from the pre-spike parameter.

Figure 10 shows the impact of different individual pre-spikes patterns. First, the situation of a single pre-spike within a paired pre-post spike association (1:1) induces no change in the pre-spike effect variable, which altered the synaptic weight variation in our STDP rule. The graphic shows that two consecutive pre-spikes coming from the visual sensory neuron, followed by the emission of a single spike from the motor neuron (2:1) (induced by the vibration neuron), modify the pre-spike effect variable in some proportion. The situation of (4:1) increases this variation. Adding other pre-spikes increases this variable but giving more importance to the first two events and less to the following one in a non-linear relation. We observed almost no effect when the pre-spike event exceeded eight consecutive spikes during the critical period. The same logic applies to the post-spike effect with slight asymmetric results (not shown). Thus, when many pre- and/or post-spikes are observed close in time (high individual spike frequencies induced by repetitive or high-amplitude stimuli) the synaptic weight variations change faster and stronger than when the spike frequencies are low. The temporal aspect of all previous pre-spikes within the temporal window is also accounted for. When the previous (2:1) situation is compared with a trial where the two consecutive spikes are more distant at the end, it offers a reduction of the pre-spike effect variable.

In the previous experiments, the external information pattern was structured and totally controlled in

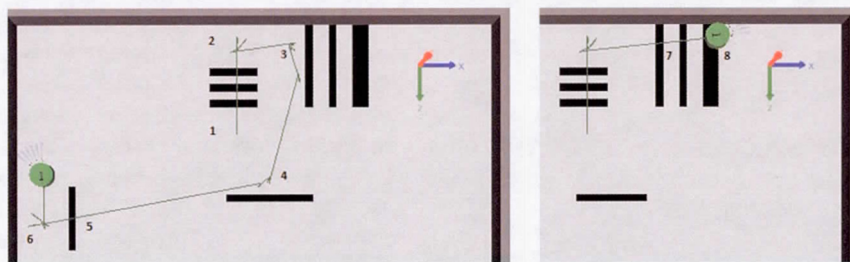


**Figure 11.** Random external flashes of light (CS) and vibration (US) inputs caught by the respective neural transducers components, leading to corresponding spikes (a–d). The dynamic integration of the induced excitatory post-synaptic potentials (PSP) is done at the associative motoneuron, leading to spike pattern (e). Continuous associations (causal and acausal) are made at the concerned synapse, causing small oscillations (y-axis scale is adjusted down 50% for observation) of the weight around the original state (h). As expected, the motoneuron spikes match the US spike pattern and lead to a cumulated reflex backward movement (i).

contrast with natural neurons, often subject to random stimuli without mentioning complex internal noise of different sources. Figure 11 illustrates the example of different such patterns of external temporal stimuli. The consecutive effect of the random sequences of pre- and post-spikes gives rise to a well-balanced variation of the synaptic change among the causal and acausal

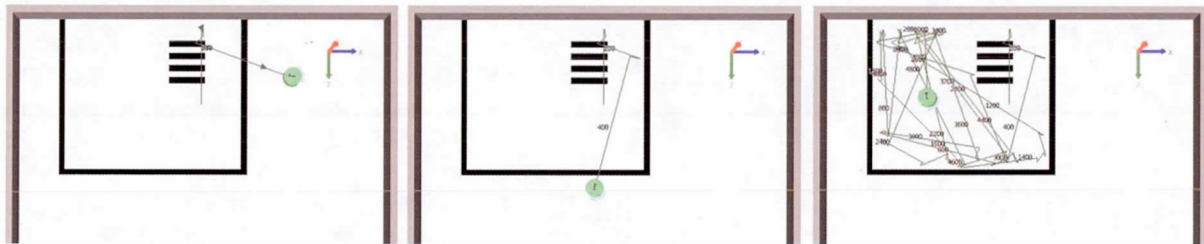
mode, the timing of the paired spike, the cumulative frequency of the paired spikes and the history of pre- and post-synaptic firing rates. Multiple trials with various stimuli frequencies show a mean located around the preferred and original steady state of the synaptic weight when the external information is unstructured. This simulation reveals no persistence in the spike pattern correlations, which do not lead to the creation of a stable association.

Among many possible scenarios, the next simulation involves a robot that learns associations between a neutral stimulus (dark strips on the floor caught by ambient light sensors) and a behavioural reflex of obstacle avoidance because of tactile sensors surrounding the robot. Considering the situations when physical robots may have undesired contact with objects, it might be wise to learn fast how to avoid these situations. Associating contextual signs in an unsupervised learning paradigm could help solve the problem. In other words, a robot able to associate and remember a CS (neutral sign) with a US (reflex avoidance behaviour) from conditioning may improve its global behaviour. In this experiment, the robot moves freely in a closed arena with walls and contextual sets of dark strips on the floor. In Figure 12 (left part), the robot's trace shows that with the neural architecture including the STDP rule, the robot does associate in one trial, the dark strips on the floor and the first contact with the wall. Consequently, the next encounter with a dark strip elicits a CR as backward and turning movement behaviour. In comparison (Figure 12 right part), the same vertical black band does not trigger any particular avoidance reflex since, in this situation, the neural architecture lacks the STDP function. In other protocols not shown, varying the darkness intensity of the



**Figure 12.** Screen shots of part of the virtual world. On the left side, the robot's artificial spiking neural network (ASNN) uses a spike-timing-dependent plasticity (STDP) rule. (1) At the start, seeing three neutral horizontal dark stripes (conditioning stimulus, CS) on the floor elicits no particular behaviour from the robot; (2) then, reaching the dark-grey delimiting wall (unconditioned stimulus, US) triggers a sequence of US → unconditioned response (UR) backward and turning movements and, at the same time, the STDP rules creates an association between the preceding CS with the US, increasing the synaptic strength of a CS → conditioned response (CR) circuit; (3) as a result, the next dark stripe seen (the first vertical bar) elicits a CR. (4) The same unsupervised learning behaviour persists for the next horizontal stripe seen. (5) Here, the robot shows a forgetting feature because the strength of the synapse was at the fringe of triggering spikes, previously stimulated by only one CS–US association (This behaviour is explaining by the backtrack sub-function include in the STDP rule). (6) Finally, the robot still exhibits an unconditioned response when facing a US. On the right figure, the same robot without the STDP rule shows no such CS → CR behaviour and sees all dark stripes as neutral stimuli (7), even when they are wider (8).





**Figure 13.** This figure shows the effect of introducing a decay factor on the persistence of previous associations in the context of the spike-timing-dependent plasticity (STDP) algorithm. On the left, the decay starts after a very short period, unlearning the causal association of the dark stripes on the floor and the collision with the wall, before encountering the next stripe. In the middle, the decay was set to a medium value, allowing the robot to react to the first dark stripe as a conditioned response but it ‘forgets’ the association when nearing the next strip. The right side of the figure shows the case when the decay factor is set to infinity, not affecting the synaptic changes made from either causal or acausal associations.

stripes modulates the behaviour in terms of more or less rapid adaptation to the context association. Also, the frequency of the pre- and post-spikes could be altered by different combinations of spatially distributed stripe patterns on the floor as well as their widths, which again modified the learning curve.

Figure 13 illustrates the effect of the decay parameter on the robot’s behaviour. When pre-post or post-pre spike pairings are made within a delimited time window, the STDP process does modulate the synaptic weight of the associative neuron in relation of the number of events occurring and the relative timing of the association. If there is no associative event context and depending on the value of the decay parameter, the STDP process starts slowly to backtrack to its original weight value. This kind of forgetting factor is not observed if the decay variable is removed from the STDP algorithm.

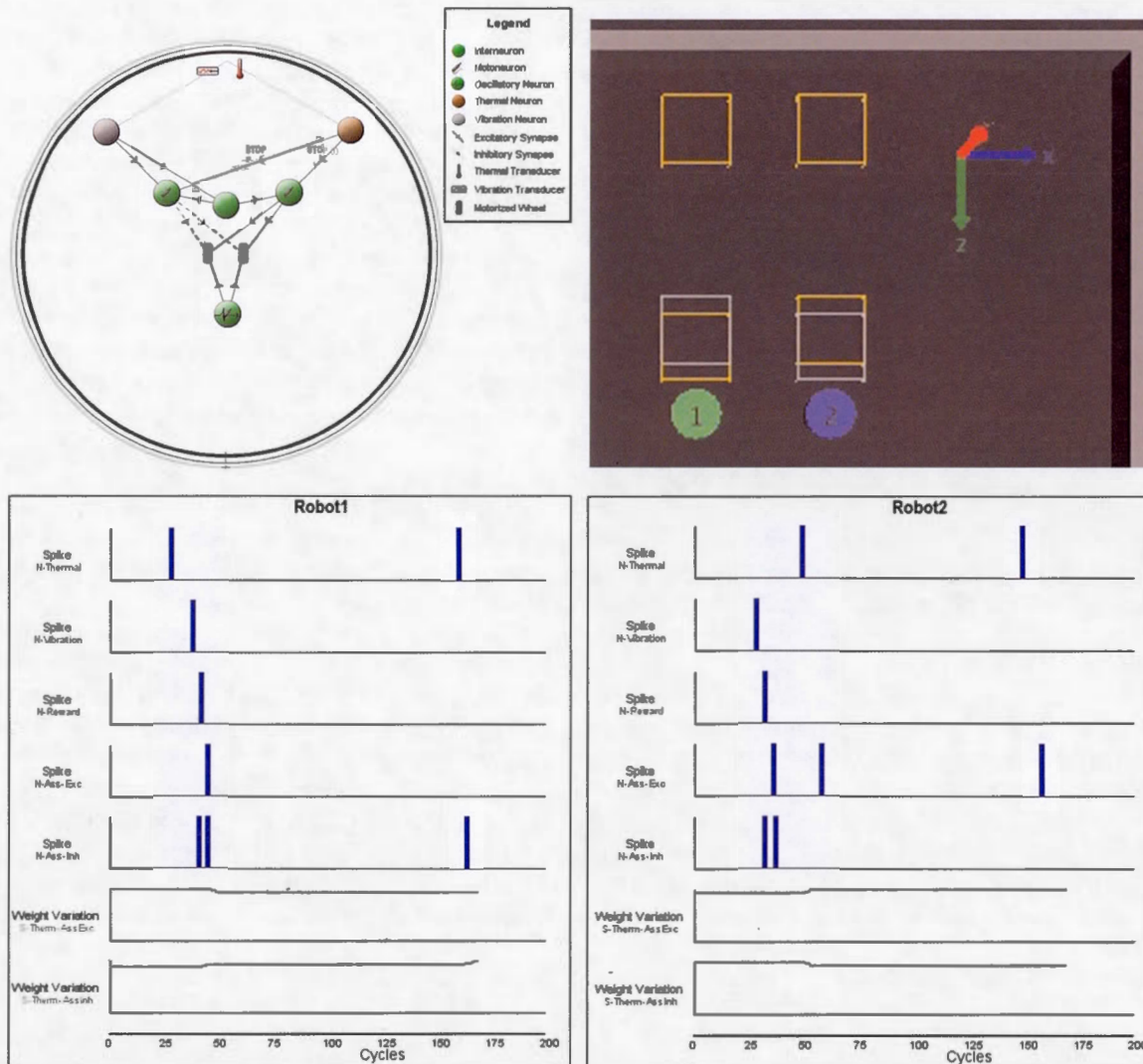
In the next simulation, we introduce the STDP rule as a potential computational means to switch between two opposing motor behaviours. This natural problem could be conceived as a bi-directional associative behaviour in biological organisms (Tanimoto, Heisenberg, & Gerber, 2004). Using the STDP diversity forms and its temporal associative features to achieve bi-directional associative behaviours may improve significantly the computational power of the learning rule in AI agents. In this respect, we present a scenario involving two virtual robots with the same neural architecture (Figure 14, top left), but confronted to two different temporal sequences of stimuli. The neural architecture enables two locomotion behaviours, one to move backward and the other to move forward, and an inverse STDP rule or STDP (i) where a pre-post spike leads to decrease the synaptic weight and a post-pre spike pattern leads to potentiation. The STDP (i) rule is only set at the synapse between the thermal neuron and the excitatory motoneuron. This biologically plausible mechanism will permit to switch from one locomotion behaviour to the other, depending of the temporal event association the robots encounter during the simulation.

The experiment proceeds as follows: robot1 encounters a CS prior to a US while robot2 encounters the reverse situation. The CS is represented by a brief thermal stimulus and the US by a single vibration wave (Figure 14, top right). The control situation follows in time, resulting in a demonstration of the learned CR. These stimuli are triggered when the robots cross certain coordinates in the map. Because the synaptic weights of the two associative neurons are set to sub-threshold firing, any paired-spikes will lead to one of the two behaviours. The bottom of Figure 14 provides all information about the precise timing of the switch.

In the last round of simulations, by performing real-time adjustments of the step cycle, we were able to reproduce the previous learning scenario with a physical modified Khepera III robot (K-Team co.), namely Tall-Boy (Road Narrows co.). For instance, in a simple conditioning task, the robot learns after a few trials to associate properly a CS (flash of light) with a US (avoidance obstacle from infrared sensors). The CR (backward movement when stimulate by a flash of light) holds for a while, and then eventually comes back to the homeostatic neutral response (not shown in the simulation). This preliminary, but convincing, incursion trial into real-world conditions confirms the validity and reproducibility of the presented data. Additional and more complete results concerning physical robots and real-world experiments are planned in the future. All virtual and physical simulations are available at the website: <http://aifuture.com/neurocode/id/13/learning-functions/associative/stdp>.

## 4 Summary and discussion

In this paper, we explored one temporal facet of conditioned learning through a modified STDP rule, from the perspective of complete AI agents. Using simulations in virtual worlds and real-world implementation, we presented a step-by step analysis of the associative mechanism in the context of a bio-inspired brain-like controller embedded in robots. From synapses to



**Figure 14.** Top left: Neural architecture producing two opposing behaviours. The left motoneuron connects to the motorized wheels with inhibitory links, while the right motoneuron does so with excitatory links. As a result, the spikes generated by these neurons will lead to a backward or forward movement (reciprocal inhibition is not included in the schema for simplicity). A contextual or reward neuron links to both motoneurons and has a synapse with the vibration neuron. Top right: This figure represents a scenario with a partial view of the virtual world. The two robots will first move forward under two consecutive stimuli pattern: robot1 starts with a thermal effect (conditioning stimulus, CS, yellow square), followed by a vibration effect (unconditioned stimulus, US, white square), while robot2 encounters the reverse situation. Afterward, both robots encounter another CS for conditioned responses (CR). Bottom left: Temporal sequence for robot1 showing the thermal neuron (CS) spikes, none followings from the excitatory motoneuron since the synaptic weight is below the firing threshold. Then, the Vibration neuron spikes (US) are followed by the contextual neuron (N-Reward), the inhibitory associative motoneuron (N-Ass-Inh) and the excitatory motoneuron (previously excited by the N-Reward) (all strong synaptic weights). As a result of a duality effect in the spike-timing-dependent plasticity (STDP) rules, the synaptic weight decreases (STDP(i): pre-post) between the Thermal neuron and the excitatory motoneuron (S-Therm-AssExc) while it increases (STDP: pre-post) between the Thermal neuron and the inhibitory motoneuron (S-Therm-AssInh). Strengthening the S-Therm-AssInh synapse provokes a spike emission from the N-Thermal when stimulated later for control, representing the predictive backward movement (CR). Note that the y-axis scales are adjusted 50% down to observe the small synaptic weights variation. Bottom right: the inverse situation happens for robot2: the US-CS sequence results in strengthening the S-Therm-AssExc synapse, leading to the opposite CR: a forward movement at the control thermal stimulus.

behaviour, the cognitive systems were designed around ASNN, a computational model of natural neural systems where the spike level resolution is the main structure of information. First, we embedded a STDP function into an ASNN, which seems to be a critical

learning component in biological counterparts. Second, instead of using a general STDP model, we designed an original algorithm that could process any spike pattern from bio-realistic features of individual and paired spike frequency parameters, a lack in most current



STDP models. Third, this ASNN was embedded in a complete AI agent, able to perceive and act within its environment. The goal was to show the benefit of adding complexity in the temporal domain from the synaptic learning rule, as validated by the observation of significant changes at the behavioural AI agent level.

We studied various temporal patterns of external stimuli implemented as CS and US onto the ASNN's synapses and the STDP learning rule. The protocol included two different input modalities involved in a specific unsupervised conditioning task. We showed that the ability of the robot to capture and integrate substantial characteristics of the information carried by the external stimuli, such as temporal order and frequency, helps the agent resolve the task successfully. It is the STDP function that allowed the embodied and situated robot to grasp the temporal complexity of these external stimuli by adapting the appropriate internal synaptic response. We showed also that a virtual instance of a robot Khepera III succeeds in performing a conditioning task as a result of the inclusion of the extended STDP function.

In any type of computational paradigm, creating an internal representation of information that reflects external real-world events within an AI agent is still a challenge. Bio-inspiration seems to be a promising approach since natural agents understand their own environment by transforming and integrating diverse external physical stimuli into a unique language of electrical modulation through different neural properties. However, mimicking lower biological functions is not a guarantee of increased intelligence in artificial artefacts. While other computational paradigms have been and are being tested to resolve real-world problems, for many researchers nature represents a suitable avenue for modelling AI agents. As such, shifting the level of abstraction from ANN to ASNN models for an AI agent's brain-controller design could be interpreted as a benefit to move toward a deeper biological plausibility and functional relevance, while increasing accuracy in the time domain. In this vein, the feasibility of incorporating STDP learning models within an ASNN is already proven, as are the advantages of inserting a non-supervised learning rule in virtual or physical AI agents. Put in different associative situations, learning from past experiences may result in observed behavioural changes. Nevertheless, combining STDP, ASNN and conditioning task in a complete AI agent framework is rarer as we mentioned previously. The motivation of this paper was thus to enhance the time-domain learning possibilities by introducing more complexity in a STDP rule, with the result of generating more flexibility in the AI agent at the behavioural level.

Paired-spikes events in two connected neurons constitute the basis of general STDP models (Zou, 2010). Synaptic plasticity implies physical changes in response to stimuli, leading to a functional perturbation of the

current state or affecting the whole efficiency of the synapse. As in the animal model, a small synaptic change sometimes results in an observable effect at the behavioural level of the whole organism. The software used in our experiments allowed us to track the entire sequence. These correlations are causal, directional and predictable, because the first cell component drives the behaviour of the next attached one, through dynamical repetitive updates and rerouting of the synaptic weight. In the natural phenomenon, these associations are thought to pass through a coincidental detector: the NMDA receptor and intermediary mediators (Shouval & Kalantzis, 2005; Sjöström & Nelson, 2002).

These events operate within a specific temporal window where past and current spike timing and frequency influence the connectivity between neurons. Under certain conditions, LTP reflects the persistence of a positive built-up of synaptic efficacy when spikes of the first neuron precede spikes of the second attached neuron. The inverse relation is also partially true with LTD. The impact is often an imbalance in the correlated processes, favouring a positive feedback, a common problem in STDP and Hebbian learning rules in general. Factors that may explain the homeostasis in real synaptic plasticity are the asymmetry and the variety of STDP processes (Gütig, Aharonov, Rotter & Sompolsky, 2003). In fact, the general heterogeneity feature of structures and functions in natural neural systems may be part of adaptive solutions in the learning domain. As a trial in this direction with the opposing behaviour experiment, the significance of using several concomitant sub-types of STDP functions instead of a unique STDP model in AI agents remains to be explored further. Including more knowledge from experimental STDP studies is also suitable to build a unified STDP model, as in the triplet problem (Gjorgieva, Clopath, Audet, & Pfister, 2012).

However, establishing that STDP is a key mechanism in natural associative learning does not resolve all the issues in AI agent bio-inspired learning. To grasp the whole picture, research, other forms of learning should be incorporated in an integrative model, including lower level mechanisms such as sensory adaptation, habituation and sensitization as non-associative functions. Also, we addressed here only one form of classical conditioning. Each of the many features of conditioning procedures merits investigations in an effort to reproduce other forms of experimental learning. A remark on the Pavlovian conditioning paradigm is that it is commonly employed in the second or minute timescale rather than in the millisecond range or step cycle observed in STDP models. To fill the gap, scientists raised the hypothesis of the so-called credit assignment or distal reward problem as does Izhikevich (2007), explaining that a slowly decaying trace of the weight change signal could be sustain from dopaminergic neurotransmitter before it disappear. Also, adding

complexity to neural properties and architectures might then prolong the association phenomenon to a longer period as with theta frequencies observed in hippocampus. It is our intention to continue studies in these promising directions to bring temporal learning in AI closer in line with real-world event association arising over multiple non-linear timescales simultaneously. Therefore, we intend to use more physical robot platforms, fully embracing the embodiment paradigm and dealing with additional challenging constraints of real-time. Also, other features of the natural STDP need to be explored, such as the implications of using different learning algorithms for inhibitory synapses implemented as heterogeneous forms of STDP rules.

The results of our experiments suggest a first-rate role for the different facets of temporality in the global intelligence phenomenon. Regarding AI agents design, we believe that one route to master the complexity of real-world information, perceived on different timescales, requires the development of well-defined supporting internal mechanisms. In an AI bio-inspired paradigm using ASNN as brain-controller, the model coherence is achieved through the representation of the outside information at the level of internal spiking neural code, where the artificial STDP function represents one inner learning mechanism. The impact of developing such a STDP function boosted with temporal features is to help the robot to understand complex environments and make appropriate adaptations as a result of learning from past experiences, processing relevant associations of various external stimuli patterns in the temporal domain. The relevant issue of this simple bio-inspired STDP learning mechanism might be generalized to make associations between more complex neural circuits and behaviours, regardless of the sensory system and the world it involved, because of the universal low-level language of the spike information.

In summary, we developed a bottom-up framework to cover some temporal aspects of a novel STDP rule. At the same time, the embedded learning function allows us to observe top-down results based on the behavioural adaptations of a virtual and physical mobile robot. This proposed STDP rule aims primarily to make progress in the field of bio-inspired learning models when put in context of ASNN serving as a robot's brain. The functional significance of using this specific STDP rule was showed through classical conditioning experiments, standing on different temporal characteristics of perceived external stimuli. Our results support the feasibility of inserting such a complex STDP process in an ASNN, with the observation of conclusive responses at the robot's behavioural level. They suggest that extended behavioural adaptation to temporal cognitive problems is possible through this STDP rule, sustaining an online synaptic plasticity process when exposed to complex and continuous external stimuli patterns that could arise simultaneously over

different timescales. To the best of our knowledge, no previous study has developed a STDP rule with deeper temporal realism, built for an ASNN controller and designed specifically for the purpose of an AI agent. We believe that this cognitive system can contribute new insights and might serve to build incrementally higher processes related to the general intelligence phenomenon.

## Notes

1. Membrane potential variation, the produced sequences = [1,13,23,30,35,38,40,41,42,43,44,46,49,55,63,100]; Parameters: Resting potential = 41; Threshold = 65; Spike = 100; Refractory period = 1; After-spike = 1; Ascending curve equation from resting state potential to Threshold:  $y = 0.1x^3 - 0.8x^2 + 2.9x + 38.8$ ; Ascending curve equation from After-spike state potential to resting state:  $y = 0.06x^3 - 1.8x^2 + 17.4x - 14.8$ .
2. This single decay parameter reflects the persistence period of a synaptic change. Currently, the factors that influence short-lasting, long or very-long lasting potentiation are not fully understood (i.e. protocol paradigm, burst, reverberation or theta rhythm in micro-circuit; Abraham, 2003; Villarreal, Do, Haddad & Derrick, 2002).

## Acknowledgements

Special thanks to Frédéric Thériault from Objectif 8 Co. for helping with programming and design issues.

## Funding

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

## References

- Abraham, W. C. (2003). How long will long-term potentiation last? *Philosophical Transactions of the Royal Society of London B*, 358, 735–744.
- Alnajjar, F., & Murase, K. (2008). A simple aplysia-like spiking neural network to generate adaptive behavior in autonomous robots. *Adaptive Behavior*, 16(5), 306–324.
- Alnajjar, F., Zin, I. B. M., & Murase, K. (2009). A hierarchical autonomous robot controller for learning and memory: Adaptation in a dynamic environment. *Adaptive Behavior*, 17(3), 179–196.
- Arena, P., De Fiore, S., Patane, L., Pollino, M., & Ventura, C. (2009). STDP-based behavior learning on TriBot robot. In *Proceedings of the SPIE*, International Society for Optical Engineering, WA, USA.
- Bi, G., & Poo, M. (1998). Activity-induced synaptic modifications in hippocampal culture: Dependence on spike timing, synaptic strength and cell type. *Journal of Neuroscience*, 18, 10464–10472.
- Bi, G., & Poo, M. (2001). Synaptic modification by correlated activity: Hebb's postulate revisited. *Annual Review of Neuroscience* 24, 139–66.
- Bohte, S. M., & Mozer, M. C. (2007). Reducing de variability of neural responses: A computational theory of spike-timing dependant plasticity. *Neural Computation*, 19, 371–403.

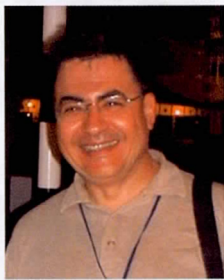


- Bohte, S. M. (2004). The evidence for neural information processing with precise spike-times: A survey. *Natural computing*, 3(2), 195–206.
- Bush, D., Philippides, A., Husbands, P., & O'Shea, M. (2010). Spike-timing dependent plasticity and the cognitive map. *Frontiers in Computational Neuroscience*, 4, 142.
- Caporale, N., & Dan, Y. (2008). Spike timing-dependent plasticity: A Hebbian learning rule. *Annual Review of Neuroscience*, 31, 25–46.
- Chiel, H. J., & Beer, R. D. (1997). The brain has a body: Adaptive behavior emerges from interactions of nervous system, body and environment. *Trends in Neurosciences*, 20, 553–557.
- Christodoulou, C., Bugmann, G., & Clarkson, T.G. (2002). A spiking neuron model: Applications and learning. *Neural Networks*, 15, 891–908.
- Cowan, W. M., Südhof, T. C., & Stevens, C. F. (2001). *Synapses*. Baltimore, MD: Johns Hopkins University Press.
- Cyr, A., Boukadoum, M., & Poirier, P. (2009). AI-SIMCOG: A simulator for spiking neurons and multiple animat's behaviours. *Neural Computation & Application*, 18(5), 431–446.
- Dan, Y., & Poo, M.-M. (2006). Spike timing-dependent plasticity: From synapse to perception. *Physiological Reviews*, 86, 1033–1048.
- Di Paolo, E. A. (2002). Spike timing dependent plasticity for evolved robots. *Adaptive Behavior*, 10, 243–263.
- Floreano, D., & Mattiussi, C. (2008). *Bio-inspired artificial intelligence: Theories, methods and technologies*. Cambridge, MA: MIT Press.
- Florian, R. V. (2010). *The chronotron: A neuron that learns to fire temporally-precise spike pattern*. Available from Nature Proceedings, <http://hdl.handle.net/10101/npre.2010.5190.3>.
- Froemke, R. C., & Dan, Y. (2002). Spike-timing-dependent synaptic modification induced by natural spike trains. *Nature*, 416, 433–438.
- Gerstner, W., & Kistler, W. (2002). *Spiking neuron models: Single neurons, populations, plasticity*. Cambridge: Cambridge University Press.
- Gjorgjieva, J., Clopath, C., Audet, J., & Pfister, J.P. (2011). A triplet spike-timing-dependent plasticity model generalizes the Bienenstock–Cooper–Munro rule to higher-order spatiotemporal correlations. *Proceedings of the National Academy of Science USA*, 108, 19383–19388.
- Gütig, R., Aharonov, R., Rotter, S., & Sompolinsky, H. (2003). Learning input correlations through nonlinear temporally asymmetric Hebbian plasticity. *The Journal of Neuroscience*, 23(9), 3697–3714.
- Izhikevich, E. M. (2003). Simple model of spiking neurons. *IEEE Transactions on Neural Networks*, 14, 1569–1572.
- Izhikevich, E. M. (2007). Solving the distal reward problem through linkage of STDP and dopamine signaling. *Cerebral Cortex Advance*, 17, 2443–2452.
- Kempler, R., Gerstner, W., & Hemmen, J. L. V. (1999). Hebbian learning and spiking neurons. *Physical Review E*, 59(4), 4498–4514.
- Kistler W. M., & van Hemmen J. L. (2000). Modeling synaptic plasticity in conjunction with the timing of pre- and postsynaptic action potentials. *Neural Computation*, 12, 385–405.
- Letzkus, J. J., Kampa, B. M., & Stuart, G. J. (2007). Does spike-timing-dependent synaptic plasticity underlie memory formation? *Proceedings of the Australian Physiological Society* 38, 15–23.
- Lieu, C., & Shapiro, J. (2007). Implementing classical conditioning with spiking neurons. In J. Marques de Sa et al. (Eds.), *ICANN 2007*, Part I, LNCS, 4668 (pp. 400–410). Berlin: Springer-Verlag.
- Maass, W., & Bishop, C. M. (Eds.). (1999). *Pulsed neural networks*. Cambridge, MA: Bradford Book, MIT Press.
- Markram, H., Lubke, J., Frotscher, M., & Sakmann, B. (1997). Regulation of synaptic efficacy by coincidence of postsynaptic APs and EPSEs. *Science*, 275, 213–215.
- Markram, H., Gerstner, W., & Sjöström, P.J. (2011). A history of spike-timing-dependent plasticity. Review article. *Frontiers in Synaptic Neuroscience*, 3(4), 1–24.
- Mehrtash, N., Jung, D., Hellmitch, H. H., Schoenauer, T., Lu, V., & Klar, H. (2003). Synaptic plasticity in spiking neural networks (sp inn): A system approach. *IEEE Transactions on Neural Networks*, 14(5), 980–992.
- Pfister, J. P. (2006). [http://biblion.epfl.ch/EPFL/theses/2006/3577/EPFL\\_TH3577.pdf](http://biblion.epfl.ch/EPFL/theses/2006/3577/EPFL_TH3577.pdf).
- Roberts, P. D., & Leen, T. K. (2010). Anti-Hebbian spike-timing-dependent plasticity and adaptive sensory processing. *Frontiers in Computational Neurosciences*, 4(156), 1–11.
- Roberts, P. D., & Bell, C. C. (2002). Spike timing dependent synaptic plasticity in biological systems. *Biological Cybernetics*, 87, 392–403.
- Rudy, J. W. (2008). *The neurobiology of learning and memory*. Sunderland, MA: Sinauer Associates, Inc.
- Schmajuk, N. A. (2010). *Mechanisms in classical conditioning*. Cambridge: Cambridge University Press.
- Shouval, H.Z., & Kalantzis, G. (2005). Stochastic properties of synaptic transmission affect the shape of spike time-dependent plasticity curves. *Journal of Neurophysiology*, 93, 1069–1073.
- Shouval H.Z., Wang S.S.-H., & Wittenberg G.M. (2010). Spike timing dependent plasticity: A consequence of more fundamental learning rules. *Frontiers in Computational Neuroscience*, 4, 19.
- Sjöström, P. J., & Nelson, S. B. (2002). Spike timing, calcium signals and synaptic plasticity. *Current Opinion in Neurobiology*, 12, 305–314.
- Song, S., Miller, K. D., & Abbott, L. F. (2000). Competitive Hebbian learning through spike-timing-dependent synaptic plasticity. *Nature Neuroscience*, 3(9), 919–926.
- Soula, H., Alwan, A., & Beslon, G. (2005). Learning at the edge of chaos: Temporal coupling of spiking neurons controller for autonomous robotic. *AAAI Spring Symposium on Developmental Robotics*, Stanford, CT.
- Tanimoto, H., Heisenberg, M., & Gerber, B. (2004). Event timing turns punishment to reward. *Nature*, 430, 983.
- Verhoog, M.B., & Mansvelder, H.D. (2011). Presynaptic ionotropic receptors controlling and modulating the rules for spike timing-dependent plasticity. *Neural Plasticity*, 2011. doi:10.1155/2011/870763.
- Villareal, D.M., Do, V., Haddad, E., & Derrick, B.E. (2002). NMDA receptor antagonists sustain LTP and spatial memory: Active processes mediate LTP decay. *Nature Neuroscience*, 5, 48–52.
- Wang, X., Hou, Z.-G., Zou, A., Tan, M., & Cheng, L. (2008). A behavior controller based on spiking neural networks for mobile robots. *Neurocomputing*, 71, 655–666.
- Zou, Q. (2009). *Computational models of spike-timing dependent plasticity: Synapses, neurons and circuits*. Germany: Lambert Academic Publishing.

### About the Authors



**André Cyr** received his B.Sc. degree in kinanthropology from the University of Quebec at Montreal in 1992, a Doctor of medicine degree in 1997 from the University of Montreal, and a family medicine specialization diploma in 1999. He is currently finishing a Ph.D. in cognitive informatics at the University of Quebec at Montreal. His thesis work addresses temporality in AI agents and, using his dual training in AI and medicine, his central interest is to investigate the general intelligence phenomenon via cognitive models embedded in robots equipped with bio-inspired neural controllers evolving through different learning paradigms.



**Mounir Boukadoum** studied Physics at the University of Algiers, Algeria, before receiving a Masters in electrical engineering degree from the Stevens Institute of Technology, USA, in 1978 and a Ph.D. degree in electrical engineering from the University of Houston, USA, in 1983. Since 1984, he has been with the University of Quebec at Montreal (UQAM), Canada, where he is now full professor of microelectronics engineering. He was director of microelectronics programs at UQAM for a 3-year mandate in 1994, 1998 and 2001, and director of the Ph.D. program in cognitive informatics from 2006 to 2009. He is currently executive member of the Microsystems Strategic Alliance of Quebec (ReSMiQ), and president of the Montreal chapter of the IEEE Computational Intelligence Society. His research interests focus primarily on intelligent instrument design and nature-inspired computing.



ANNEXE D

ARTICLE 3 : CONDITIONNEMENT OPÉRANT

FRONTIERS IN NEUROROBOTICS





# Operant conditioning: a minimal components requirement in artificial spiking neurons designed for bio-inspired robot's controller

André Cyr\*, Mounir Boukadoum\* and Frédéric Thériault\*

Computer Science Department, Cognitive and Computer science, Université du Québec à Montréal, Montréal, QC, Canada

## Edited by:

Marc-Oliver Gewaltig, Ecole Polytechnique Fédérale de Lausanne, Switzerland

## Reviewed by:

Frank Van Der Velde, University of Twente, Netherlands  
Berat Denizduran, Ecole Polytechnique Fédérale de Lausanne, Switzerland

## \*Correspondence:

André Cyr, Mounir Boukadoum and Frédéric Thériault, Cognitive and Computer Science, Université du Québec à Montréal, 201 Avenue du Président-Kennedy, H2X 3Y7, Montréal, Québec, Canada  
e-mail: andré.cyr1@videotron.ca; boukadoum.mounir@uqam.ca; ftheriault@cvm.qc.ca

In this paper, we investigate the operant conditioning (OC) learning process within a bio-inspired paradigm, using artificial spiking neural networks (ASNN) to act as robot brain controllers. In biological agents, OC results in behavioral changes learned from the consequences of previous actions, based on progressive prediction adjustment from rewarding or punishing signals. In a neurobotics context, virtual and physical autonomous robots may benefit from a similar learning skill when facing unknown and unsupervised environments. In this work, we demonstrate that a simple invariant micro-circuit can sustain OC in multiple learning scenarios. The motivation for this new OC implementation model stems from the relatively complex alternatives that have been described in the computational literature and recent advances in neurobiology. Our elementary kernel includes only a few crucial neurons, synaptic links and originally from the integration of habituation and spike-timing dependent plasticity as learning rules. Using several tasks of incremental complexity, our results show that a minimal neural component set is sufficient to realize many OC procedures. Hence, with the proposed OC module, designing learning tasks with an ASNN and a bio-inspired robot context leads to simpler neural architectures for achieving complex behaviors.

**Keywords:** operant conditioning, robot, learning, spiking neurons, adaptive behavior, bio-inspired agents

## INTRODUCTION

Learning is well recognized by the scientific community as a major feature of intelligence. In biological agents, this concept implies behavioral adaptations from experiences, and it is essential for confronting the high variability of the real world. In this context, learning brings real-time flexible solutions that a reactive response mode alone cannot afford. As an evolutionary process during an organism's lifespan, learning allows an anticipation of the near future by doing actions differently. Many description levels of learning exist as partial explanations of the empirical data founded on the phenomenon. Among alternatives, the neurobiological cellular level seems a coherent approach to understand any learning processes deeply. Whether studies on learning refer to natural or artificial agents, extensive research exists in both domains. In artificial intelligence (AI), including the robotic field, learning rules flourish and the direction depends on the contextual paradigm and goals they are utilized for Watkins (1989) and Sutton and Barto (1998).

Among the existing approaches to robot controllers, the bio-inspiration trend (Bekey, 2005; Floreano and Mattiussi, 2008; Pfeifer et al., 2012) suggests that the principles underlying natural intelligence, of which the learning phenomenon, can be understood theoretically and may be used as inspiration to realize the core computation in AI agents. The derived computational models, once embedded in robotic "brains" with sensory inputs and motor outputs, may then emulate natural intelligent behavior.

One important candidate for the development of bio-inspired controllers is the artificial neural network (ANN), an abstraction of the natural counterpart. Like real neurons, artificial neural units use stored weights (memory) to perform the integration of one or more inputs at an input node and generate an activation value. The weights are set via a learning strategy and a transmission function uses the integration result to generate the neural output, typically action potentials that feed one or more other neural inputs in the ANN. The ANN has many interesting properties: the ability to process implicit and fragmentary data, high computation parallelism, and the capacities of generalization, discrimination, classification and pattern completion, all possible in a variety of application contexts.

Among several declinations of ANNs, the latest is represented by the artificial spiking neural network (ASNN) sub-family (Maass, 1997; Gerstner and Kistler, 2002). ASNNs distinguish themselves from preceding ANN models in that the underlying algorithm is inherently suited to resolve temporal cognitive problems. Instead of using rate coding to model the individual neural firing, the time stamping used by the ASNN allows timing, synchronization, and asymmetrical temporal correlations (Izhikevich, 2003) between spikes. ASNNs are also well-adapted for real world applications (see Ponulak and Kasiński, 2011, for a survey). In particular, they do not require prior batch training as much as ANNs and their weight adjustment process is typically much faster than the offline gradient descent algorithm that is

often used for ANNs (Wilson and Martinez, 2003). For instance, an ASNN with unsupervised learning rules may learn solutions in just one trial from a first correlated paired of spikes, hence exhibiting a critical ability in a real-time context, particularly for organisms with short life spans, which cannot afford a learning mechanism based on slow synaptic adjustments. Finally, a desirable feature that ASNNs share with ANNs is the possibility of add-on learning rules to cope with uncertainty and complexity when handling dynamic information.

In this work, we use a dynamic ASNN adapted for virtual and physical robotic simulations. It relies on a discrete-time leaky integrate-and-fire (LIF) neural model, implemented as a program that uses lookup tables to specify the synaptic and neural properties. A LIF neuron is a mathematical abstraction of the biological neuron that puts emphasis on the integrative and output functions parts. Simple LIF models (Gerstner and Kistler, 2002; Izhikevich, 2003) consider that time-varying post-synaptic potentials (PSP) are the sources of inputs currents to the neuron (positive or negative), with a subsequent dynamic effect on the membrane potential. The process is typically modeled with the following state equation:

$$C \frac{du(t)}{dt} = -\frac{1}{R}u(t) + \left( i_0(t) + \sum_n w_{ij}(t) \right)$$

where  $u(t)$  stands for the cell membrane potential,  $C$  for the membrane capacitance,  $R$  for its resistance, and  $i$  represent the input currents coming through synapses (weights)  $w$ . A spike is generated whenever  $u(t)$  reaches a predetermined (and constant) threshold  $\theta$ , after which the membrane potential is reset to a resting value  $\theta_0$  for a refractory period  $T_r$ . The leaky term  $\frac{1}{R}u(t)$  allows emulation of passive ionic diffusion, a characteristic forcing the membrane potential to come back slowly to its original equilibrium state.

The integration of the input presynaptic spikes to form the input currents could take different curve shapes, but it is generally approximated by an exponential function to represent the membrane potential's ability to reach the firing threshold. Hence, an input current is typically represented by:

$$i(t) = \int_0^8 s(\tau - t) e^{-\frac{t}{\tau_s}}$$

where  $s()$  represents the presynaptic spike train and  $\tau_s$  is the integration time constant.

It should be clear that the previous equations are a simplification of the actual organization and neurophysiologic processes of a real neuron. In particular, the spatial structure, the mapping of pre-synaptic to post-synaptic potentials and the role of neurotransmitters are lumped into simple synaptic weights. However, the model is sufficient to account for many useful neural phenomena.

Our LIF neural model includes all the previous elements, but we use lookup tables to represent the dynamics of the various inputs variables, and ASNNs can be built using them thanks to a software called SIMCOG that we developed (Cyr et al., 2009) (see Supplementary material for details and starting parameters).

Here, we use it to demonstrate operant conditioning (OC) in a bio-inspired robotic paradigm, when applied in different contexts through an ASNN.

Obviously, reproducing the natural OC learning process must stand on critical, permanent, and tractable neural components. In this respect, we suggest a constant corpus of elements that are common in multiple OC scenarios. First, we establish a simple model of a habituation rule, a non-associative learning process. Habituation is defined by many temporal features (Thompson and Spencer, 1966; Rankin et al., 2009), but the main ones are a decreasing response to persistent stimuli and a recovery part when they cease. It involves more-or-less persistent modifications at the pre-synaptic element. The functional impact of habituation is thought to minimize redundant information, filtering input and enhancing stimuli novelty (Marsland et al., 2005). The justification to include this learning rule in our OC model is to avoid reflex behaviors from constant input (McSweeney et al., 1996), acting similarly to an intrinsic action selection mechanism. To this end, the original ASNN of SIMCOG was recently enhanced with a novel computational model of habituation (Cyr and Boukadoum, 2013), extended for temporal features.

The next step in a neurobiological perspective about realizing an OC model is the implementation of classical conditioning (CC) (Pavlov, 1927). This type of learning allows agents, through association between repetitive pairs of conditioned (neutral) and unconditioned (meaningful) stimuli to eventually produce a conditioned response that is the same as the unconditioned one (Schmajuk, 2010). CC is considered a passive associative learning skill, since agents have no power to intervene in the conditioning process. For instance, the sea-slug *Aplysia Californica*, one well-known lower neural system animal model, exhibits aversive CC of the gill and siphon withdrawal reflex (Hawkins, 1984; Glanzman, 1995). One plausible biologically cellular mechanism that partially explains event association in CC is the spike-timing dependent plasticity (STDP) process (Bi and Poo, 1998; Markram et al., 2011), which consists mainly of a directional and causal modification of synaptic strength in relation to the precise timing of paired-spikes within a specific temporal window. As such, an STDP computational model of associative learning was also recently proposed in the ASNN of SIMCOG (Cyr and Boukadoum, 2012).

OC also consists in a form of associative learning where agents modify their behaviors by associating stimuli and responses, actively changing their behavior from past acquired experiences, by the motivation of rewards or punishments. Hence, agents alter their normal behaviors from acquired knowledge by "evaluating" the acts they previously made and the current changes. Pioneers (Thorndike, 1911; Skinner, 1938; Hull, 1943) left many researches as legacy, establishing the foundation of the subject. OC is a phenomenon recognized across modalities and developmental phases (Valente et al., 2012) in almost all natural neural species. In particular, the study of the OC mechanism has greatly benefited from invertebrate models such as the sea slug *Aplysia* (Brembs, 2003). Still, the structural elements and the precise dynamics involved in the process remain to be determined. As a starting point in the quest for a minimal components requirement defining OC, the present literature in psychology (Frieman,



2002; Chance, 2009) generally underscores three high-level steps: an agent must perform a behavior; a modification must follow this action (the phenomenon); the consequence of the modification must reinforce the action positively or negatively. As a result, the change in the action, situation, or environment can either strengthens or weakens the behavior. The change could also affect the form, range, frequency, persistence, timing, and magnitude of the behavioral response. In addition, the OC procedures generally includes recurrent cycles of spontaneous, voluntary or random actions followed by a reinforcement stimulus, whether it is removed or presented, rewarding or punishing and resulting in higher or lower reproduction of the contributive behavior. Finally, the behaviors could happen in chain and in a non-stationary environment, which dramatically increases the temporal order and complexity of the phenomenon (Touretzky and Saksida, 1997).

It has already been demonstrated that OC requires a different internal molecular mechanism than CC (Lorenzetti et al., 2006). Nevertheless, the latter seems to be involved in OC (Holland, 1993), reflecting the basic terminology of convergence and divergence where CC refers to stimuli-reinforcer (CS-US) and OC to a response-reinforcer (Behavior-Reinforcer) contingency. A comparative study (Baxter and Byrne, 2006) of these associative processes indicates that the convergence of the neural activity and the reinforcer in OC lead to a change in the neural membrane properties (modification of the input resistance and burst threshold), while CC is concerned with the modulation of a tagged synaptic site (Izhikevich, 2007). Also, several lines of evidence point to the dopamine as the excitatory neurotransmitter involved in the OC reinforcement process (Bédécarrats et al., 2013). Researchers in the domain are currently exploring some interesting molecular tracks (Lorenzetti et al., 2008) at the resolution of an analog single cell model (Brembs et al., 2002) to confirm which and how the components serves as the coincidental detector.

Computationally modeling the OC process at different conceptual levels depends on the desired outcome. The idea of developing an OC model using a rate-coded ANN is not new (Grossberg, 1971), nor is the idea to embed this learning process in a virtual robot (Graham et al., 1994) or a real one (Gaudiano and Chang, 1997). These OC models share a common goal: Apply the comprehension of the natural learning mechanism for the benefit of physical robots. For example, Gaudiano AND Chang demonstrate how real robots could learn from a modified version of Grossberg's CC and OC model in an unknown environment, a simple navigation task of avoiding or approaching from rewarding or punishing cues. These opposing behaviors match different sensory cues. Perhaps, one limitation of that article is that a robot could eventually learn approaching walls or any objects to receive a reward or, inversely, avoid light when punished. Still, based on rated-coded ANNs, their model sustains the generalization of environments and minimizes the tuning parameters for an ego-centric robotic framework. Since timing of events is an important factor in OC procedures, ASNN is preferred.

On the other hand, attempts to develop a time-coded ASNN as basis for an OC model embedded in a bio-inspired robotic platform is more recent (Arena et al., 2009; Helgadóttir et al., 2013; Soltoggio et al., 2013) and merits further development. Our

contribution consists in a simple computational OC model, built with an elementary micro-circuit that can be generalized and applied in multiples OC scenarios, demonstrating that this learning process could involve only a small number of specific cellular elements. More precisely, an input feeding Cue neuron, an Action neuron and a Predictor neuron receiving a reward or punishment seem sufficient to compose the core of OC when connected by specific directed links and few learning rules such as habituation and STDP.

We base our approach on the fact that the OC phenomenon is seen in the *Aplysia*, and also in smaller neural organisms such as *C. elegans* (300 fixed neurons), which are already known for non-associative and associative learning (Qin and Wheeler, 2007) for adaptive behaviors. Moreover, this flat worm is perceived to also exhibit a derived dopamine implication in many behavioral modulation roles (Vidal-Gadea and Pierce-Shimomura, 2012). As such, the low neural count in these organisms and their multiple adaptive learning capabilities may imply that a simple integrative OC model may be sufficient to reproduce them. However, until now, few computational models are integrating such bio-inspired neural knowledge for robotic training purposes and general AI behavioral applications. In this work, we embed an OC model in an ASNN acting as a robot brain, using learning functions such as habituation and STDP, and a few neurons and synaptic links. To reach the goal, the paper includes several OC procedures of incremental level of complexity, using four general categories of contextual OC: (1) Increased behavior with the expectation of receiving a positive reinforcer; (2) Increased behavior with the expectation of not receiving a negative reinforcer; (3) Decreased behavior with the expectation of not receiving a positive reinforcer; (4) Decreased behavior with the expectation of receiving a negative reinforcer.

Scenario A uses a simple contextual procedure to demonstrate the minimal components required to build an invariant OC kernel. It corresponds to the first item in the list above (bold). Scenario B switches to the opposite OC primitive polarity within the same learning experiment; in few time-steps online, the robot learns from the same sensory cue to avoid doing the action depending of the contextual type of reinforcer. In scenario C, we suggest a more elaborate and dynamical experiment by varying the parameters of the minimal OC components requirement. In scenario D, we embed a representative ASNN implementing an OC procedure in a real robot platform. In all four scenarios, the OC kernel remains unchanged.

To summarize, despite the fact that the exact mechanism of OC is still unknown, we propose in line with the current cellular level understanding of the natural process a simple OC model for virtual and physical robots by embedding a neural core within an ASNN framework. As novelty, we suggest that OC applied in general contexts could be understood in terms of an invariant minimal component requirement and show the benefit of integrating the habituation and the STDP learning rules into the kernel. Our intention is not to provide new theories of the natural phenomenon nor revising extensive features of OC. The scope of this OC model is limited to simple OC contexts, and remains to be tested in more complex situations. Nevertheless, we above all target generalizable functional outcomes in robots to further



ground this important adaptive learning process into real world applications, beside natural intelligences.

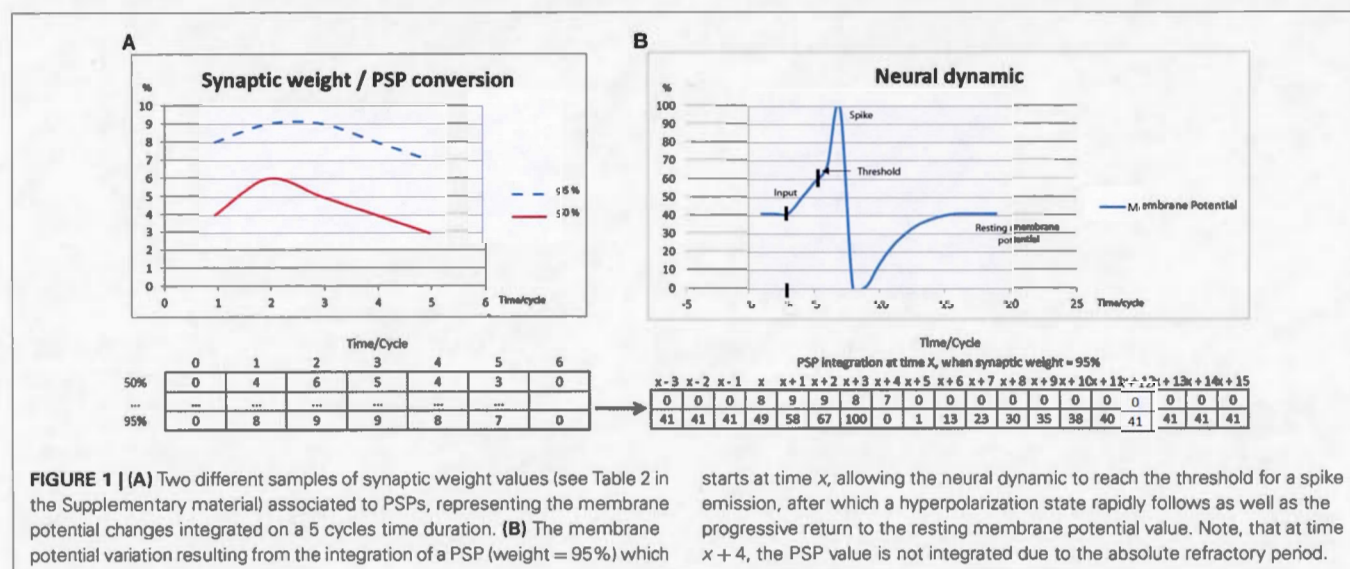
The balance of the paper is structured as follows: In the next section, methodology, we describe the dynamics of the ASNN and the software used. The various tasks, scenarios and results are detailed in Section 3. Finally, a discussion and future work close this article.

## METHODOLOGY

The OC procedures of this paper were implemented with the SIMCOG software, but other environments could have been used. In order to create a generic simulation environment, all the variables in SIMCOG related to the ASNN dynamics are normalized in the interval [0–100] and only integer values are used. With reference to **Figure 1**, a neuron is represented by an activation state whose membrane potential amplitude varies between 0 and 100, where 0 stands for the maximum of hyperpolarization, 63 is the threshold for firing and 100 the level of an emitted spike. The synapse between two neurons is represented by an excitatory or inhibitory link whose weight is a value varying also between 0 and 100. The different weights serve as indices in a vector table containing the sampled data of normalized PSP curves. Depending on the input and the current membrane potential state of a neural unit, a non-linear integration is performed across time, as done in a biological neuron, leading to a new membrane potential value and eventually triggering a spike emission if the firing threshold is reached. One feature of the ASNN editor in SIMCOG is the possibility to add sensory transducers to neurons. These unique components are intended to simulate the natural elements in the realm of robot sensors. Hence, these transducers form a bridge between physical or virtual world stimuli and their neural attachment, creating a sensory stimulus or graded receptor potential that serves as input to the ASNN. The sensory inputs also use a [0–100] scale, pointing to vectors that correspond to variations of membrane potential to receive. The same logic is applied for the motor neuron outputs, emitting sequences of current toward actuators, producing actions in virtual or physical world.

Lookup tables represent a strategy in computational modeling, including ASNN models (Ros et al., 2006; Alhawarat et al., 2013). Their use is motivated by the fact that neural models and variables are not always described by functions in closed form and, sometimes, experimental curves are the only precise data to rely on. In our work, we also used it to account for the lower CPU capacities in the robot's model. Notice that, since the curve shapes are similar for graded receptor potentials, PSPs and graded motor potentials, the same lookup tables is used (see Table 2 in the Supplementary material). A common fixed length of 5 cycles and a maximum change of 10% maintain coherence in the system regarding equivalent values in mV and ms of input/output and empirical results in the literature. Thus, the values that fill the vector table are set qualitatively, based on a progressive, non-linear stimulus-intensity scale. All values are integers in the range [0–100].

For all the neural branching parts (transducer/synapse/actuator) in the ASNN, different additive learning rules could be applied to obtain a desired modulation control (see Supplementary material, Equations 3–5). A sensory adaptation rule is used to restrict the data flow at the transducer site. An exponential decay factor is applied for a reduction percentage to occur on a short time period, thus allowing the model to deal with brief as well as constant stimuli. A habituation rule that acts between two neurons plays the same role essentially, but it filters redundant input information on longer and more variable temporal scales depending on the intensity, magnitude, frequency, and inter-stimulus interval of the input pattern. The habituation rule is a key factor in the OC dynamic we present, regulating the spikes traffic in the process of neural integration. Depending of the temporal input pattern, different exponential decay factors modulate the synaptic weights as well as recovery functions when the stimuli cease. Finally, we also applied a STDP learning rule at the synaptic site, for modification of the synaptic weight in relation of the precise pre-post spike timing (i.e., on a short time scale window, pre-post spike occurrences strengthen the link between two neurons and post-pre spike occurrences weaken it. All these



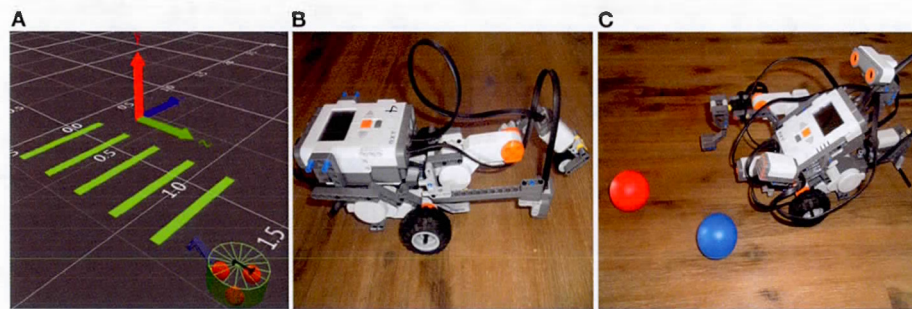


learning rules are also implemented with integer values through lookup tables and built qualitatively from standard curve shapes in the literature.

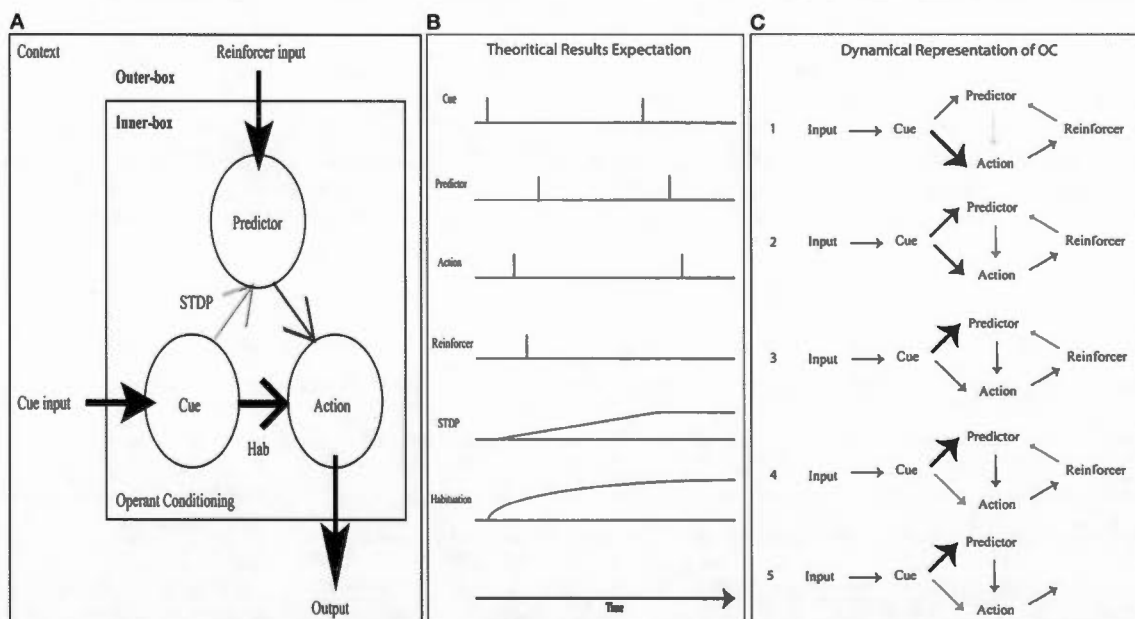
In the virtual experiments, we choose a 15 cm diameter frame, two motorized-wheels at the back for locomotion and a support wheel at the front for balance (**Figure 2A**). A two-segment arm at the front of the virtual robot allows pressure on the ground such as color circles into complex scenarios. The physical robot is built using a Lego Mindstorms NXT 2.0 and intends to mimic the virtual one (**Figures 2B,C**).

In the simplest OC procedure, the ASNN includes a kernel (inner-box in **Figure 3A**) built with three neurons, three

synapses, one habituation rule, and one STDP synaptic rule. To accommodate the virtual and physical robots with a wide range of OC scenarios, the kernel is completed with a contextual outer-box that provides input from at least two transducers, output to one actuator, and three additional synapses that connect to the inner-box kernel (outer-box in **Figure 3A**). These last synapses come from the reinforcer input, the cue input and to the action output. We submit that this inner-box and outer-box combination has the minimum number of components to achieve OC procedures in varying contexts for a bio-inspired robotic paradigm (see **Figures 3B,C** for alternative dynamical representations of the OC model).



**FIGURE 2 | (A)** Representative third person view of NeuroSim, the 3D-world simulator in the SIMCOG suite, showing a picture of the virtual robot when attempting to cross several green bars on the floor. **(B,C)** Similar structures with physical robots (Lego Mindstorms NXT 2.0).



**FIGURE 3 | OC simulation with a minimal neural component set. (A)** The inner-box consists of a Cue neuron feeding an Action neuron with a strong synaptic link (bold) and a Predictor neuron with a weak link (pale); the link between Predictor and Action is of intermediary strength. A habituation rule between Cue and Action allows decreasing the associated synaptic weight with time, leading to fade the Action. When a reinforcer input triggers a spike from Predictor, consecutively to a spike from Cue, the associated synaptic link

increases due to STDP learning. It eventually allows Cue alone to stimulate Predictor and from this new path Action neuron. In the contextual outer-box, the input and output can be abstracted to any type of neural structure. **(B)** Dynamic representation of OC with the crucial components. Using a positive reinforcer (reward), the results are simulated through a timeline: Predictor leads to an Action without any Reward. **(C)** Dynamic step representation of the OC learning phenomenon with an emphasis on synaptic strength changes.

The contingencies of the OC procedures are tributary of the array of robot behaviors and the contextual environment, regardless of virtual and physical robot or natural agents. Still, the spontaneous production of behaviors is a necessary condition for any OC procedures. In all the virtual experiments that we conducted, the robots evolve in a closed 3D-world that contains different color bars painted on the floor, tactile sensors to detect the robot's arm pressing the floor, surrounding sound sensors to detect sound emissions from the robots and, vibration/light/heat emitters for cue input or rewarding/punishing reinforcers. Depending of the protocol, the robot's actions consist of emitting sound, moving, pressing the floor with an attached arm or pushing color bricks. Thus, the robot has restricted navigation skills and the *a priori* knowledge comes from the behavioral patterns inherited from the ASNN. The subsequent modulation of the responses comes uniquely from the learning rules, and the results depend on the temporal associations of the sensory cues, the reinforcers and the actions, which change with context.

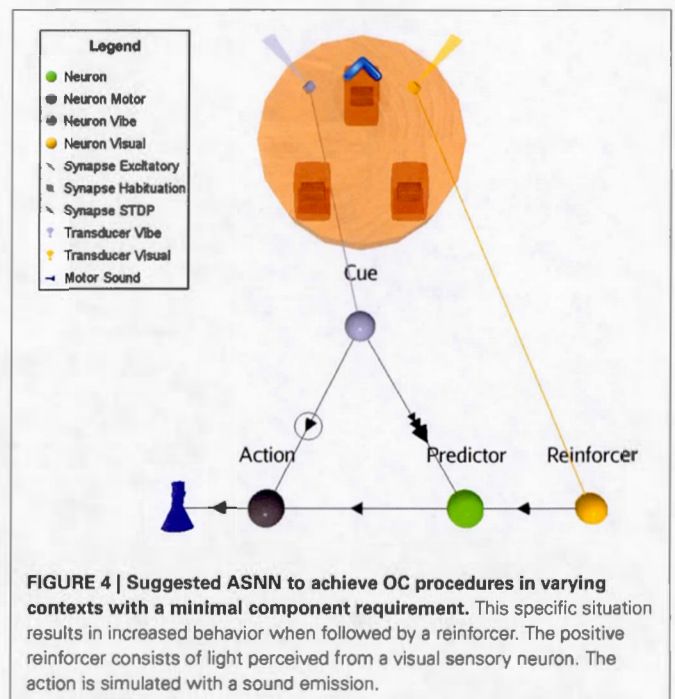
### TASKS, SCENARIOS, AND RESULTS

Pecking a light or a color disk is a common task with pigeons in many learning studies. Pressing levers into a skinner-box or performing maze escapes are also stereotype experiments with rats, producing a vast literature in the learning domain. These apparently simple behaviors are more complex than they appear, and many variables manipulations are needed to observe the data on CC and OC procedures. In animal learning, the positive reinforcer usually consists in giving water or food when the animal is hungry or thirsty. In contrast, the negative reinforcer is often represented by brief electrical shocks, air puffs, or direct touches. The associative neutral stimuli used in natural contexts include flashes of light, sound, odor, color disks and mechanical devices (button, lever, and maze).

In this work, we abstract the behavior with sound emission in scenarios A and B. In scenario C, the action is to press on the floor and in scenario D to push blocks. The willingness to do or not an action can be resolved in many ways. Reflex action is a simple approach that we adopt here, though intrinsic neural properties could also lead to spontaneous endogenous random or periodic actions. Otherwise, any central pattern generators, value systems, or action selection mechanisms could also result in actions. In all experiments here, we explicitly propose an external reinforce, though again, alternative experiments could embed internal reinforcers such as a second ASNN standing for an intentional or motivational behavior. Scenarios A and B involve a positive reinforcer implemented as flashes of light and a negative reinforcer implemented as a vibration wave. In all scenarios, manual parameter adjustments were done. Efficiency was not a priority and more formal tuning methods could be used, such as genetic algorithms or other ANNs. Additional information on the parameter values used for scenario C is provided in the Supplementary material.

#### SCENARIO A: PRIMITIVE OC PROCEDURE (INCREASED BEHAVIOR/POSITIVE REINFORCER)

The ASNN associated with the robot configuration is shown in **Figure 4**. It allows realizing the primitive OC procedure of

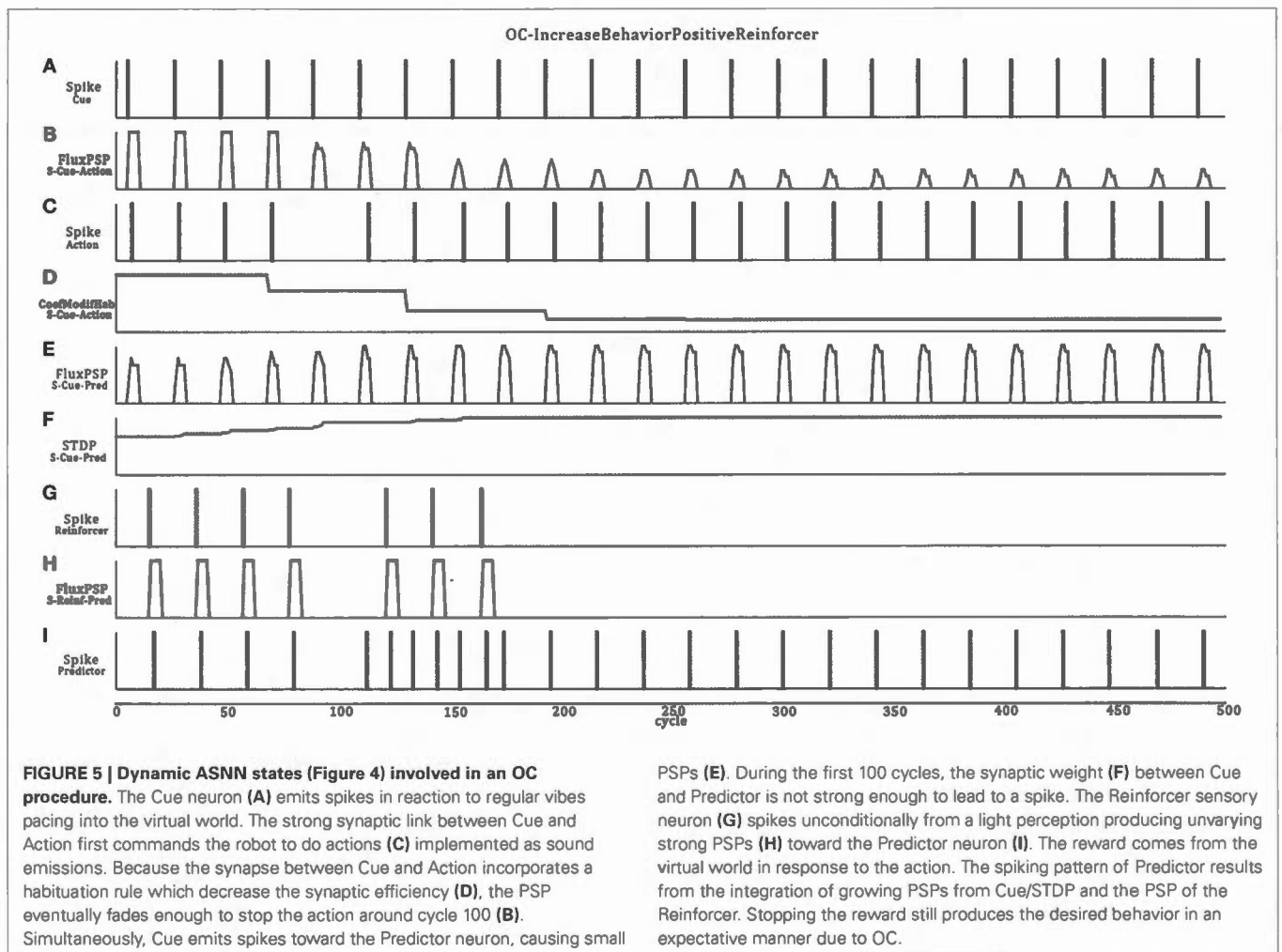


increasing (all or none in this experiment) the behavior when followed with a positive reinforcer. This basic configuration can be generalized to the other primitive cases and it represents the proposed minimal neural components requirement. In particular, swapping the excitatory synaptic link for an inhibitory one between the Action neuron and its actuator will result in decreasing the behavior. The attribution of a positive or negative aspect of the reinforcer corresponds to an internal value, which depends on the context. Hence, the four primitive OC scenarios could be embedded in the same structure.

Referencing to **Figure 5**, the Cue neuron reads periodic vibrations emitted from the virtual world. Consequently, Cue emits spikes to the Action neuron that triggers a reflex sound emission from the robot. Flashes of light (Reinforcer) are emitted into and from the virtual world (supervised procedure) consecutively to the robot's action for a certain period. In this ASNN configuration, the endpoint is to provoke a sound emission in the expectation of a rewarding light. At the synaptic link between Cue and Predictor, the STDP learning rule associates the sound emission (from the Action neuron) and the ensuing positive reinforcer (reward input) to the Predictor neuron. Discontinuing the reward (around cycle 175), reveals the ascending strength of this synaptic link, leading to predict alone the reward from the cue.

Four control situations can exist (not graphically shown): (1) Without the habituation rule, the action never stops and removes the benefit of predicting the reward, as well as producing the action without the possibility to stop it in the case of a negative reinforcer; (2) Without the STDP rule, the reward is always requested to produce the action and the cue will never be strong enough to predict the reward (3) If the reward precedes the action, the STDP rule will decrease the synaptic weight and the action stops rapidly, highlighting the non-predictive value of the link (4) Finally, stopping the cue will lead to "forget" the association, an





extinction feature of OC procedures. We also did simulations on all learning rule permutations within the OC core and found no other possibility to produce OC.

The synaptic weights must be well balanced in order for the OC to work properly. Within the previous learning rules permutation tests, we randomized several times the initial synaptic weights and found no other concluding experiments. More precisely, we achieved the OC procedure only when there was a high synaptic weight between Cue and Action, and a low weight between Cue and Predictor.

#### SCENARIO B: TWO OPPOSITE PRIMITIVE OC PROCEDURES (INCREASE BEHAVIOR/POSITIVE REINFORCER) THEN (DECREASE BEHAVIOR/NEGATIVE REINFORCER)

This scenario represents a small increment in the ASNN complexity (Figure 6). We kept the simple structure of scenario A, but add the possibility of punishment (negative reinforcer) with a brief heat wave. One can see the duplication of the minimal component structure. Note that these imbricates neural circuits could be used in separate OC procedures. In comparison to scenario A, some interneurons are added to link both circuits and explicitly highlight the mutual exclusive opposition of behaviors (sound on/off)

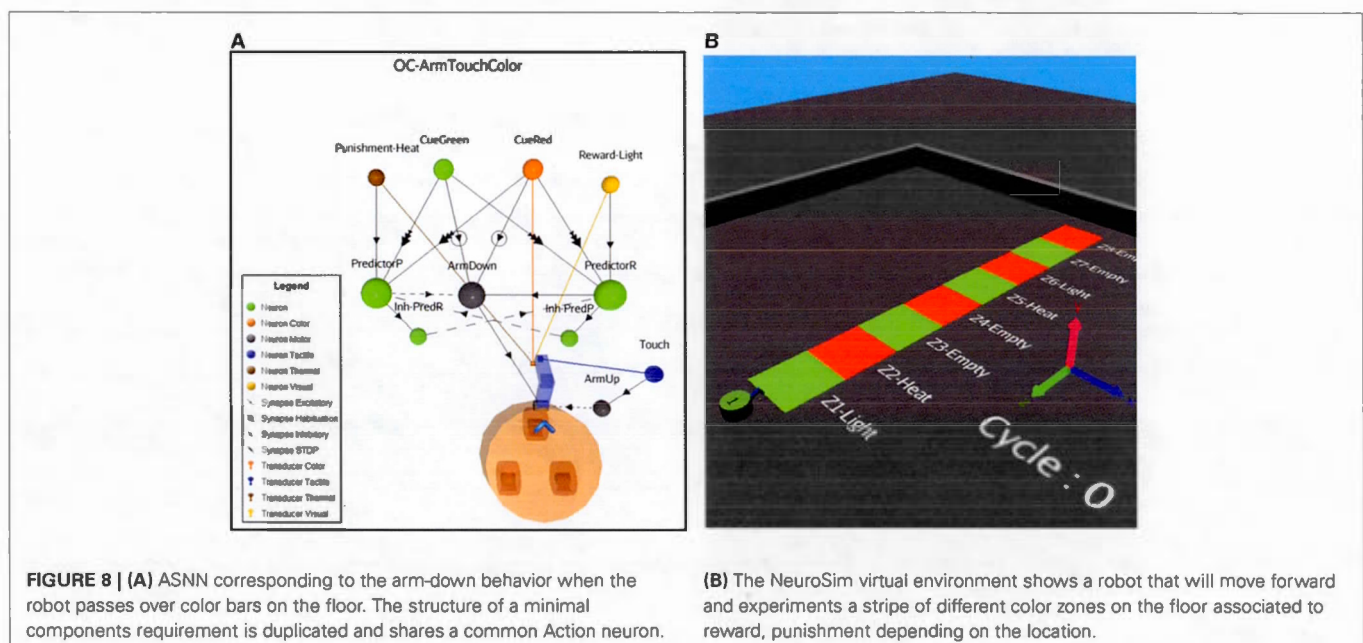
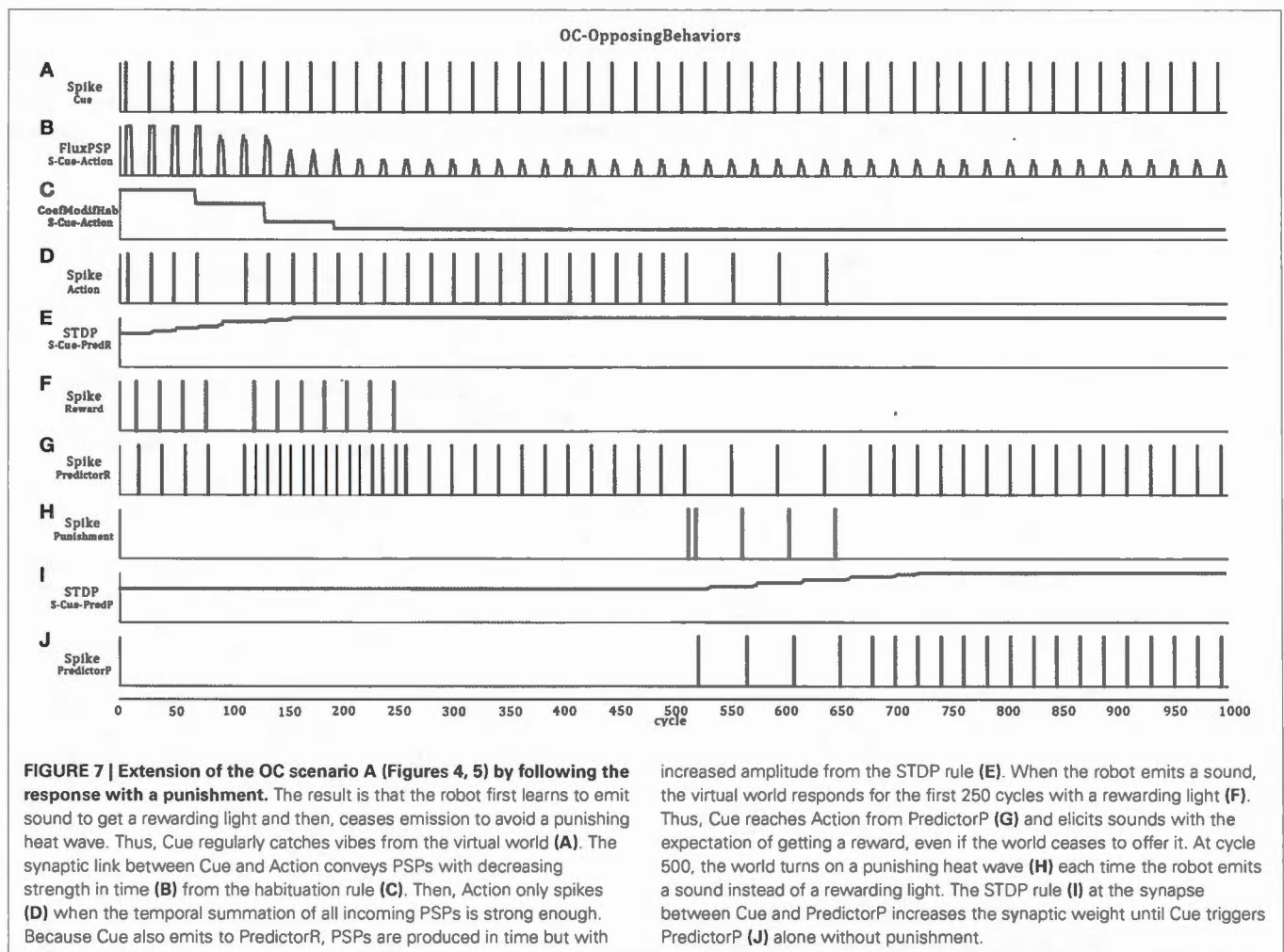
and the effect of the reinforcer (heat/light). The neural dynamic of this OC procedure (Figure 7) leads to reverse the behavior of doing a sound to get the rewarding light by stopping the action from learning by punishment. This ASNN could be abstracted to any other actions like escape behaviors. One can see the persistence of the PredictorR spike pattern from the Cue input. In fact, the link is reinforced each time Cue spikes. When ceased (not shown), the STDP eventually enters in a recovery mode and simply forgets the association with a previous reward. The same logic could be applied with PredictorP.

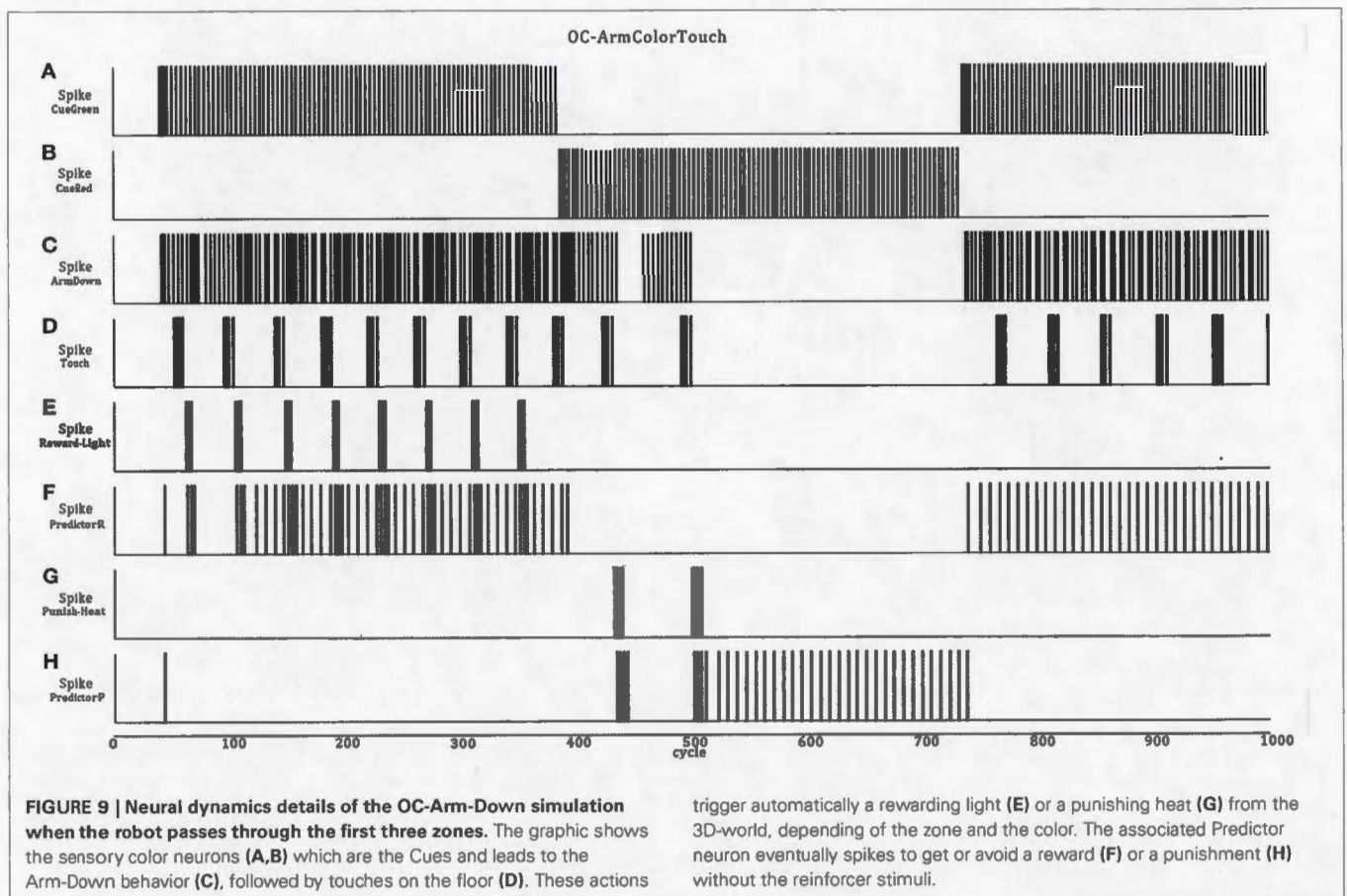
#### SCENARIO C: PRESSING COLOR X IN EXPECTATION OF REWARDS OR PUNISHMENTS

In this more complex scenario, we keep the OC kernel untouched, but replace some external elements and add others. In Figure 8A, instead of a reflex sound response when the robot perceives a vibration stimulus, two new Cue neurons are introduced, now sensitive to color patches on the floor. Cue-Green is a sensory neuron responding to the green color and Cue-Red to the red color. Also, to encompass a larger type of input patterns, including constant stimuli, we add a fast adaptation learning rule (negative exponential decay factor with a fast recovery parameter) to









elements or copies of the minimal kernel requirement are all that is needed for the OC phenomenon to proceed.

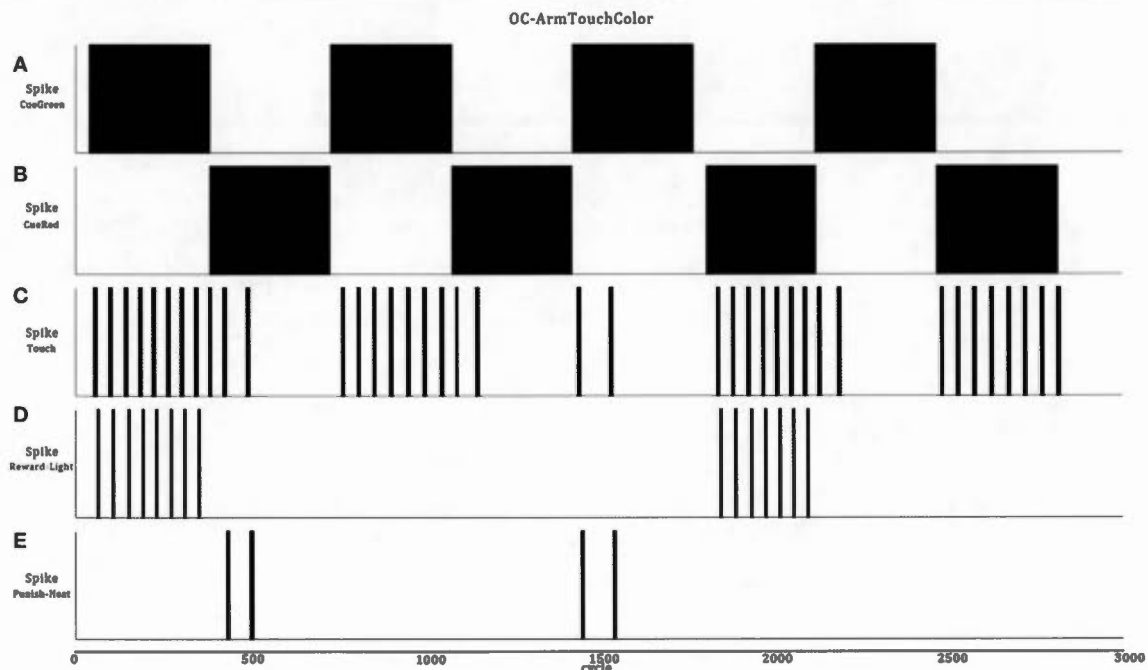
The challenge in low-level explanations of the OC mechanism is to find the convergence of behavioral consequences and reinforcer. We have proposed a simple neural circuit pointing to clear elements at the cellular resolution level. The OC model seems general enough to maintain explanations in many contexts.

The power of OC learning is, in a given context, to shift dynamically a neutral spontaneous behavior toward a preferred behavior from the association between one action and its consequence. The context is often coming from sensory cues. The value of the outcomes must also be embedded somehow and should be part of the dynamical process. Persistence, recall, and reversibility are also essential plastic learning features that lead to learning adaptation. In scenarios B and C, we showed these modulated capacities from changing opposite behaviors online and from reversing the neutral type of initial cues. The conclusion is that a stimulus, a behavior or a consequence could be good one day and bad another day. Thus, a computational model of OC should be concerned by all these essential characteristics.

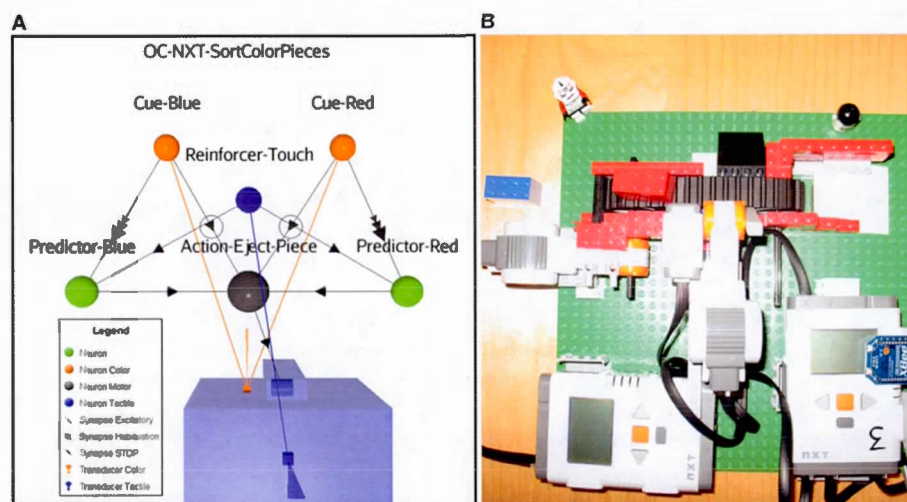
In scenario D, we demonstrated that the OC process was possible in a physical robot by transferring the ASN into it. We also varied the neural architecture as well as the temporal aspects of some parameters. By randomizing the pieces order

(Cues), we showed that the STDP rule could modulate the synaptic weights with different time intervals, as long as the ISI (inter spike interval) is included in the time-window of the STDP rule. Also, the presentation of the piece-cue-stimulus was time-varying, depending of the user that feeds the chain. Moreover, because the timing of the touch (Reinforcer) was also realized with a user (imprecise), it brings some other interesting temporal variation into the OC procedure. In a robotic perspective, sorting color items is easy with programming. Changing the color on the fly is also possible, but the system should previously have programmed the feature. With our OC model, a supervised procedure could reverse any color pieces (or any other object's features). We showed also that no prior knowledge is needed because all cues, actions, and reinforcer are considered neutral and could be changed at any time. Such flexibility could be useful in several applications in the AI domain.

In a general OC goal-directed action model, the Predictor acts by pointing the desired behavior among sets of actions by previously associating a given behavior to a given context that imply rewarding or punishing consequences. The transition of circumstantial action into a habit for expectancies represents one feature of the OC phenomenon. We find that this OC feature, as an experience-dependant plasticity process, seems realizable with a small ASN which includes few, but crucial minimal components requirement.



**FIGURE 10 |** Integral path of the robot when it passes over eight color zones (A,B). The data are the same as in Figure 9 but, the second half portion reverses the associated reward (D) or punishment (E) stimuli in regard of the color bars. At the end, the robot learns to avoid pressing the green bars and presses the red one (C).



**FIGURE 11 |** (A) ASNN of a robot sorting color pieces; it chooses between a red or blue one depending on an external touch reinforcer. (B) Conveyor belt with bricks and robot (No. 3) flipping from a half rotation movement with a

side arm to eject the pieces. A color sensor located at the front of the robot allows it to read the colors of the pieces. Also, a touch sensor is attached for external reinforcing.

As criticisms, the proposed OC including CC, habituation, and sensory adaptation models needs some examples (arbitrary) of stimuli-spike-action for the demonstration. We have not exhaustively explored the boundaries, nor defined the exact relationship between all the temporal parameters of the learning rules that are merged into the OC model. Thus, the limit values of the cue, action, and reinforcer remain to be tested

further in terms of strength, rate, and ISI parameters to complete the computational model, even though it was not our primary goal to develop a formal OC model. Also, value systems should be implemented in the OC model (Krichmar and Röhrbein, 2013); particularly when considering that learning drives do not always come as external reinforcers (Santucci et al., 2013).



In this paper, it was convenient to apply the reinforcer in the same temporal windows as the STDP, habituation, and sensory adaptation learning rules. The distal problem solving, credit assignment or simply put, how to link an input stimuli to a delayed reinforcer are highlighted by many researchers (Sutton and Barto, 1998; Izhikevich, 2007) and different theoretical models support the idea at the cellular level with as example, the tagging synapse hypothesis (Päpper et al., 2011). Moreover, in the fruit flies heat-box, OC procedures (Brembs, 2011) show that they learned in few minutes to avoid a specific punishing tube. Thus, longer temporal considerations manner and should be included for a complete OC model.

In a next model, we plan to include a neural trace component with a decay factor in association with a general feedback system. Extending the temporal window of the coincidence would then increase the “good” synaptic link. It was not a specific aim of this article to implement distant temporal gap within the OC process, though, to our knowledge, we don’t see much obstacle foreseeing this temporal characteristic.

Also, introducing a dopamine-like reinforcer as a main excitatory input for rewarding signal seems a logical step inclusion in a future OC model if achieving higher cognitive processes is a target (Wang et al., 2011). Implementing a particular neuromodulator in an ASNN could translate into a different shape of the PSP with a higher magnitude peak and a shorter time window, offering similar dynamic features of the natural component. In the same line, adding cholinergic inhibitory inputs could increase the ASNN modulation possibilities as in biological conditions (Chubykin et al., 2013).

The current understanding of the OC process seeks the convergence of behavior and reward signal at the molecular level through the synergistic interaction between different types of mediators (i.e., calcium ion and calmodulin-sensitive adenylyl cyclase concentrations). At the end, the coincidence detector may be understood by only one molecule, inducing changes in the neural membrane properties, which produce CC and OC as distinctive associative learning processes (Brembs and Plendl, 2008). We clearly don’t want to go that far in our present computational OC model though it is still an option. We prefer to focus on the functional and temporal aspects of the learning process for bio-inspired robotic purposes and concrete AI applications.

As extension of this work, several other behavioral alternatives should be implemented to validate this simple OC model in complex situations. As such, we intend to challenge this learning concept in decision-making property (Nargeot and Simmers, 2011) and its possible implication in the action selection mechanism. Adding chaotic feature to the ASNN through differential equations instead of lookup tables may find benefit in the richness of the behavioral response. Also, more complex robots should embed the OC model to appreciate the real scope of this computational learning rule. Overall, with this OC core model, building incremental complex dynamical scenarios taking advantage of habituation, STDP, and OC intrinsic characteristics may offer more than additive behavioral adaptation in neurorobotics applications.

Also, studying non-elemental forms of learning is far more difficult and characterize higher cognition in animals (Giurfa,

2007). Exploring shaping behaviors, collective-decision in heterogeneous cognitive abilities, negative pattern discrimination [learn to discriminate a binary compound stimulus and reinforce A and/or B but not AB, (A+, B+, AB-)], feature neutral discrimination (B+, AC+, AB-, C-) or mastering transitive inference rules as in  $A > B$ ,  $B > C$  then  $A > C$  base on OC and CC processes may be a path to follow.

In conclusion, this article showed a computational OC model underscoring a minimal component requirement in terms of specific cellular elements as explanation of the learning process. This original OC model was presented in line with the current understanding of the neuroscience knowledge at the cellular level of comprehension of the mechanism. We applied this OC model in a specific bio-inspired robotic paradigm through an incremental, but simple level of complexity within the proposed scenarios. The true impact and limitations of this OC model remains to be determined in a wider spectrum of applications. Yet, the singular modularity of the minimal component requirement certainly opens the door to resolve interesting tasks by AI agents from this unique structure in the OC learning process.

## SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <http://www.frontiersin.org/Journal/10.3389/fnbot.2014.00021/abstract>

## REFERENCES

- Adiwatari, M., Nazih, W., and Eldesouki, M. (2013). Studying a chaotic spiking neural model. *Int. J. Artif. Intell. Appl.* 4, 107–119. doi: 10.5121/ijaiia.2013.4508
- Arena, P., Fortuna, L., Frasca, M., and Patané, L. (2009). Learning anticipation via spiking networks: application to navigation control. *IEEE Trans. Neural. Netw.* 20, 202–216. doi: 10.1109/TNN.2008.2005134
- Baxter, D. A., and Byrne, J. H. (2006). Feeding behavior of Aplysia: a model system for comparing cellular mechanism of classical and operant conditioning. *Learn. Mem.* 13, 669–680. doi: 10.1101/lm.339206
- Bédécarrats, A., Cornet, C., Simmers, J., and Nargeot, R. (2013). Implication of dopaminergic modulation in operant reward learning and the induction of compulsive-like feeding behavior in Aplysia. *Learn. Mem.* 20, 318–327. doi: 10.1101/lm.029140.112
- Bekey, G. A. (2005). *Autonomous Robots: From Biological Inspiration to Implementation and Control*. Cambridge, MA: MIT Press.
- Bi, G. C., and Poo, M. M. (1998). Synaptic modifications in cultured hippocampal neurons: dependence on spike timing, synaptic strength, and postsynaptic cell type. *J. Neurosci.* 18, 10464–10472.
- Brembs, B. (2003). Operant conditioning in invertebrates. *Curr. Opin. Neurobiol.* 13, 710–717. doi: 10.1016/j.conb.2003.10.002
- Brembs, B. (2011). Spontaneous decisions and operant conditioning in fruit flies. *Behav. Processes* 87, 157–164. doi: 10.1016/j.beproc.2011.02.005
- Brembs, B., Lorenzetti, F. D., Reyes, F. D., Baxter, D. A., and Byrne, J. H. (2002). Operant reward learning in Aplysia: neuronal correlates and mechanisms. *Science* 296, 1706–1709. doi: 10.1126/science.1069434
- Brembs, B., and Plendl, W. (2008). Double dissociation of PKC and AC manipulations on operant and classical learning in *Drosophila*. *Curr. Biol.* 18, 1168–1171. doi: 10.1016/j.cub.2008.07.041
- Chance, P. (2009). “Chapter 5,” in *Learning and Behavior*, 6th Edn., (Belmont, CA: Wadsworth), 119–174.
- Chubykin, A. A., Roach, E. B., Bear, M. F., and Hussain Shuler, M. G. (2013). A cholinergic mechanism for reward timing within primary visual cortex. *Neuron* 77, 723–735. doi: 10.1016/j.neuron.2012.12.039
- Cyr, A., and Boukadoum, M. (2012). Classical conditioning in different temporal constraints: an STDP learning rule for robots controlled by spiking neural networks. *Adapt. Behav.* 20, 257–272. doi: 10.1177/1059712312442231



- Cyr, A., and Boukadoum, M. (2013). Habituation: a non-associative learning rule design for spiking neurons and an autonomous mobile robots implementation. *Bioinspir. Biomim.* 8:016007. doi: 10.1088/1748-3182/8/1/016007
- Cyr, A., Boukadoum, M., and Poirier, P. (2009). AI-SIMCOG: a simulator for spiking neurons and multiple animat's behaviours. *Neural Comput. Appl.* 18, 431–446. doi: 10.1007/s00521-009-0254-2
- Floreano, D., and Mattiussi, C. (2008). *Bio-Inspired Artificial Intelligence: Theories, Methods and Technologies*. Cambridge, MA: MIT Press.
- Frieman, J. (2002). *Learning and Adaptive Behavior*. Belmont, CA: Wadsworth.
- Gaudio, P., and Chang, C. (1997). "Adaptive obstacle avoidance with a neural network for operant conditioning: experiments with real robots," in *IEEE, Proceedings of CIRA* (Monterey, CA).
- Gerstner, W., and Kistler, W. (2002). *Spiking Neuron Models: Single Neurons, Populations, Plasticity*. Cambridge, UK: Cambridge University Press. doi: 10.1017/CBO9780511815706
- Giurfa, M. (2007). "Invertebrate cognition: nonelemental learning beyond simple conditioning," in *Invertebrate Neurobiology, Chapter 12*, eds G. North and R. J. Greenspan (San Diego, CA: Cold Spring Harbour Laboratory Press), 281–308.
- Glanzman, D. L. (1995). The cellular basis of classical conditioning in *Aplysia Californica* – it's less simple than you think. *Trends. Neurosci.* 18, 30–36. doi: 10.1016/0166-2236(95)93947-V
- Graham, J., Alloway, T., and Krames, L. (1994). Sniffy, the virtual rat: simulated operant conditioning. *Behav. Res. Met. Instrum. Comput.* 26, 134–141. doi: 10.3758/BF03204606
- Grossberg, S. (1971). On the dynamics of operant conditioning. *J. Theor. Biol.* 33, 225–255. doi: 10.1016/0022-5193(71)90064-6
- Hawkins, R. D. (1984). A cellular mechanism of classical conditioning in *Aplysia*. *J. Exp. Biol.* 112, 133–128.
- Helgadóttir, L. I., Haernicke, J., Landgraf, T., Rojas, R., and Nawrot, M. P. (2013). "Conditioned behaviour in a robot controlled by a spiking neural network," in *Proceeding NER2013* (San Diego, CA), 891–894.
- Holland, P. C. (1993). Cognitive aspects of classical conditioning. *Curr. Opin. Neurobiol.* 3, 230–236. doi: 10.1016/0959-4388(93)90215-K
- Hull, C. L. (1943). *Principles of Behavior*. New York, NY: Appleton-Century-Crofts.
- Izhikevich, E. M. (2003). Simple model of spiking neurons. *IEEE Trans. Neural Netw.* 14, 1569–1572. doi: 10.1109/TNN.2003.820440
- Izhikevich, E. M. (2007). Solving the distal reward problem through linkage of STDP and dopamine signaling. *Cereb. Cortex* 17, 2443–2452. doi: 10.1093/cercor/bhl152
- Krichmar, J. L., and Röhrbein, F. (2013). Value and reward based learning in neurobots. *Front. Neurobot.* 7:13. doi: 10.3389/fnbot.2013.00013
- Lorenzetti, F. D., Baxter, D. A., and Byrne, J. H. (2008). Molecular mechanisms underlying a cellular analog of operant reward learning. *Neuron* 59, 815–828. doi: 10.1016/j.neuron.2008.07.019
- Lorenzetti, F. D., Mozzachiodi, R., Baxter, D. A., and Byrne, J. H. (2006). Classical and operant conditioning differentially modify the intrinsic properties of an identified neuron. *Nat. Neurosci.* 9, 17–19. doi: 10.1038/nn1593
- Maass, W. (1997). Networks of spiking neurons: the third generation of neural network models. *Neural. Netw.* 10, 1659–1671. doi: 10.1016/S0893-6080(97)00011-7
- Markram, H., Gerstner, W., and Sjöström, P. J. (2011). A history of spike-timing-dependent plasticity. *Front. Synaptic Neurosci.* 3:4. doi: 10.3389/fnsyn.2011.00004
- Marsland, S., Nehmzow, U., and Shapiro, J. (2005). On-line novelty detection for autonomous mobile robots. *Rob. Auton. Syst.* 51, 191–206. doi: 10.1016/j.robot.2004.10.006
- McSweeney, F. K., Hinson, J. M., and Cannon, C. B. (1996). Sensitization-habituation may occur during operant conditioning. *Psychol. Bull.* 120, 256–271. doi: 10.1037/0033-2909.120.2.256
- Nargeot, R., and Simmers, J. (2011). Neural mechanisms of operant conditioning and learning-induced behavioral plasticity in *Aplysia*. *Cell. Mol. Life Sci.* 68, 803–816. doi: 10.1007/s00018-010-0570-9
- Päpper, M., Kempter, R., and Leibold, C. (2011). Synaptic tagging, evaluation of memories, and the distal reward problem. *Learn. Mem.* 18, 58–70. doi: 10.1101/lm.2024811
- Pavlov, I. P. (1927). *Conditioned Reflexes*. Oxford: Oxford University Press.
- Pfeifer, R., Lungarella, M., and Iida, F. (2012). The challenges ahead for bio-inspired "soft" robotics. *Commun. ACM* 55, 76–87. doi: 10.1145/2366316.2366335
- Ponulak, F., and Kasinski, A. (2011). Introduction to spiking neural networks: information processing, learning and applications. *Acta Neurobiol. Exp.* 71, 409–433.
- Qin, J., and Wheeler, A. R. (2007). Maze exploration and learning in *C. elegans*. *Lab Chip* 7, 186–192. doi: 10.1039/b613414a
- Rankin, C. H., Abrams, T., Barry, R. J., Bhatnagar, S., Clayton, D. F., Colombo, J., et al. (2009). Habituation revisited: an updated and revised description of the behavioral characteristics of habituation. *Neurobiol. Learn. Mem.* 92, 135–138. doi: 10.1016/j.nlm.2008.09.012
- Ros, E., Carrillo, R., Ortigosa, E. M., Barbour, B., and Agis, R. (2006). Event-driven simulation scheme for spiking neural networks using lookup tables to characterize neuronal dynamics. *Neural Comput.* 18, 2959–2993. doi: 10.1162/neco.2006.18.12.2959
- Santucci, V. G., Baldassare, G., and Mirolli, M. (2013). Which is the best intrinsic motivation signal for learning multiple skills? *Front. Neurobot.* 7:22. doi: 10.3389/fnbot.2013.00022
- Schmajuk, N. A. (2010). *Mechanisms in Classical Conditioning*. Cambridge: Cambridge University Press. doi: 10.1017/CBO9780511711831
- Skinner, B. F. (1938). *The Behavior of Organisms: An Experimental Analysis*. New York, NY: Appleton-Century-Crofts.
- Soltoggio, A., Lemme, A., Reinhart, F., and Steil, J. J. (2013). Rare neural correlations implement robotic conditioning with delayed reward and disturbances. *Front. Neurobot.* 7:6. doi: 10.3389/fnbot.2013.00006
- Sutton, R. S., and Barto, A. G. (1998). *Reinforcement Learning: An Introduction*. Cambridge, MA: MIT Press.
- Thompson, R. F., and Spencer, W. A. (1966). Habituation: a model phenomenon for the study of neuronal substrates of behavior. *Psychol. Rev.* 73, 16–43. doi: 10.1037/h0022681
- Thorndike, E. L. (1911). *Animal Intelligence: Experimental Studies*. New York, NY: Macmillan.
- Touretzky, D. S., and Saksida, L. M. (1997). Operant conditioning in skinnerbots. *Adapt. Behav.* 5, 219–247. doi: 10.1177/105971239700500302
- Valente, A., Hung, K. H., Portugues, R., and Engert, F. (2012). Ontogeny of classical and operant learning behaviors in zebra fish. *Learn. Mem.* 19, 170–177. doi: 10.1101/lm.025668.112
- Vidal-Gadea, A. G., and Pierce-Shimomura, J. T. (2012). Conserved role of dopamine in the modulation of behavior. *Commun. Integr. Biol.* 5, 440–447. doi: 10.4161/cib.20978
- Wang, L. P., Li, F., Wang, D., Xie, K., Wang, D., Shen, X., et al. (2011). NMDA receptors in dopaminergic neurons are crucial for habit learning. *Neuron* 72, 1055–1066. doi: 10.1016/j.neuron.2011.10.019
- Watkins, C. J. C. H. (1989). *Learning from Delayed Rewards*. Ph.D. thesis, Cambridge University, Cambridge.
- Wilson, D. R., and Martinez, T. R. (2003). The general inefficiency of batch training for gradient descent learning. *Neural. Netw.* 16, 1429–1451. doi: 10.1016/S0893-6080(03)00138-2

**Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 25 October 2013; accepted: 02 July 2014; published online: 25 July 2014.  
 Citation: Cyr A, Boukadoum M and Thériault F (2014) Operant conditioning: a minimal components requirement in artificial spiking neurons designed for bio-inspired robot's controller. *Front. Neurobot.* 8:21. doi: 10.3389/fnbot.2014.00021  
 This article was submitted to the journal *Frontiers in Neurobotics*.  
 Copyright © 2014 Cyr, Boukadoum and Thériault. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

ANNEXE E

ARTICLE 4 : SÉLECTION D'ACTIONS

JOURNAL OF ROBOTICS

## Research Article

# Action Selection and Operant Conditioning: A Neurorobotic Implementation

**André Cyr and Frédéric Thériault**

*Département d'Informatique, Université du Québec à Montréal (UQAM), Succursale Centre-Ville, Case Postale 8888, Montreal, QC, Canada H3C 3P8*

Correspondence should be addressed to André Cyr; andre.cyr1@videotron.ca

Received 23 January 2015; Revised 7 April 2015; Accepted 14 May 2015

Academic Editor: Maki K. Habib

Copyright © 2015 A. Cyr and F. Thériault. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Action selection (AS) is thought to represent the mechanism involved by natural agents when deciding what should be the next move or action. Is there a functional elementary core sustaining this cognitive process? Could we reproduce the mechanism with an artificial agent and more specifically in a neurorobotic paradigm? Unsupervised autonomous robots may require a decision-making skill to evolve in the real world and the bioinspired approach is the avenue explored through this paper. We propose simulating an AS process by using a small spiking neural network (SNN) as the lower neural organisms, in order to control virtual and physical robots. We base our AS process on a simple central pattern generator (CPG), decision neurons, sensory neurons, and motor neurons as the main circuit components. As novelty, this study targets a specific operant conditioning (OC) context which is relevant in an AS process; choices do influence future sensory feedback. Using a simple adaptive scenario, we show the complementarity interaction of both phenomena. We also suggest that this AS kernel could be a fast track model to efficiently design complex SNN which include a growing number of input stimuli and motor outputs. Our results demonstrate that merging AS and OC brings flexibility to the behavior in generic dynamical situations.

## 1. Introduction

The vast topic of action selection (AS), including decision-making, behavioral choice, and behavior-switch as nomenclatures, is thoroughly explored from different perspectives of comprehension, levels of resolution, and scientific communities [1]. The AS biological phenomenon results from a neural process that leads to the observation of an agent doing one action over several others. The precise neural substrate underpinning this mechanism is not yet discovered [2, 3]. Even though many insights [4, 5] point toward how to simulate the AS natural process in artificial agents, there is no consensus to approach this cognitive phenomenon. In this view, the neurorobotic domain aims to study AS from bioinspirations but applied for artificial intelligence (AI) and robotics purposes [6]. As a premise, building controllers for unsupervised autonomous robots necessarily requires a dedicated mechanism to operate behavioral transitions. Moreover, in real world, these actions should be adaptive instead of being ruled by fixed-patterns. Thus, a simulated

AS process should be flexible enough to cope with changing environments. These adaptive behaviors could come from learning functions which act as modulators of the AS process.

Artificial spiking neural networks (SNN) [7] have been successfully used as brain-controllers for robots, and several researches have proposed different computational models implementing AS through this specific experimental paradigm [8, 9]. A major aspect of SNN is to understand the information process at the level of a single spike [10]. Therefore, timing of spikes can be used for temporal event correlations and associative learning. As such, it could be interesting to study an AS mechanism in combination with an operant conditioning (OC) process, since we anticipate that these processes add more flexibility to switch behavior from their interactions, sharing both the ability of specifying actions.

The function of an AS process is to decide between different actions depending on the context. As a matter of fact, invertebrate neural organisms like *C. elegans* [11, 12], cnidarians [13], and fruit flies [14] do well in choosing among several



actions with only small circuits of neurons. Those include command-center neurons and central pattern generators (CPG) [15–18], which are well recognized for their intrinsic oscillation property. A deduction that could be drawn is that modeling an AS process does not necessarily require complexity as in high brain structures. Thus, the working hypothesis for the emulation of an AS process states that a simple mechanism should then be derived. In this paper, we use a basic CPG neural structure which helps in simulating an AS process containing sensory inputs, motor outputs, and decision neurons [19].

We propose to study the AS process within a SNN framework, targeting bioinspired robots controllers. Our first motivation is to combine AS and OC processes in a single neuro-robotic model. The main goal is to build a simple yet adaptive AS mechanism merged with the plasticity feature of an OC learning rule, while both occur under a dynamical scenario. A second objective is to develop a fast track method for implementing general AS processes into SNN. This research was driven by the fact that it is still a challenge to create a robot controller with the ability to learn from multiple sensory cues and actions in a SNN paradigm.

*Theoretical Background.* In neurosciences, the drive to accomplish a behavior emerges from a real-time dynamic of external sensory cues and internal values, where the different competitive neural signals ultimately orient the agent toward one preferred action. In a psychological view of the AS problem, serial processes occur from sensors to motors ending in a behavioral choice. According to the literature in computational cognitive science, the affordance competition hypothesis [20] argues that such a process is parallel and implies a prior specification of possible actions from ongoing sensory inputs. Specifically, when dynamical processes include several feedback loops in high neural structures and an attentional mechanism, the brain focuses on a specific winning action while continuously searching for other actions to do, depending on the context [21].

In the robotic domain, computational models of the AS process have been proposed (stochastic accumulator, linear-ballistic accumulator, and integrated accumulator models) [22, 23] as well as CPG in conjunction with SNN [24]. Since only a few studies in that field have investigated the AS process using SNN as bioinspired brain-controllers, our study takes another step in this direction. Therefore, our focus is on the close interaction between AS and the OC learning function, which we propose as a novelty in the domain.

Empirically, the angle of the AS problem was often to reach an optimal solution with a statistical approach [5, 25] or reproducing biological data. In our research, we wanted to consider the modulation factors that may influence the dynamic of an AS model from its interaction with a learning rule. In this perspective, a learning skill may improve a robot's choice of actions to determine the future solutions. OC consists in one of these primary learning functions allowing cognitive agents to associate a feedback from their own actions. The natural OC process is well understood at the level of invertebrates [26]. Therefore, among others [27, 28], learning with OC represents one potential modifier of the AS

mechanism, perhaps allowing more flexibility from synaptic plasticity features in their adaptive behaviors. From its own past actions, which gave rewarding or punishing feedbacks, a robot may eventually pick up a different action, accelerating or decelerating the bias toward an oriented alternative.

We address these questions of the AS process combined with OC by evaluating a simple scenario in virtual and physical robots. This current work does not focus on extensive tests nor evaluates the overall computational impact of the parameters involved in the AS-OC models. It was also beyond the scope of this paper to challenge other AS approaches. Despite these limitations, we show a biologically plausible core base of these mechanisms in a neuro-robotic implementation. A benefit for robots to include the AS and OC critical processes is undeniable, since most physical robots are now able to perform a rich selection of actions that may be organized in hierarchical priorities, sequential fixed-patterns, competitive actions, and conflicting parallel behaviors.

In resume, we show an AS mechanism based on a CPG structure and few elementary neural units. This AS process was subject to modulation when merged with an OC learning rule. Together, these processes offer more flexibility to choose the best action under dynamical and variable contexts. Further demonstrations in more complex scenarios remain to be studied.

## 2. Methodology

*2.1. The Spiking Neural Model.* We propose a simple scenario to explore the AS and OC interrelation, explaining both processes in a neuro-robotic paradigm. The robot's controller consists in artificial neural units connected by synapses. Our SNN model [29], similar to standard leaky integrate-and-fire neuron models, is based on a membrane potential variation, integrating nonlinearly, and temporally ongoing inputs through the SNN (1). In these neurons, when the membrane potential reaches a specific threshold, an all-or-none action potential is triggered. To start the CPG dynamic at the beginning of a simulation, a realistic neural property of endogenous pacemaker is implemented from adding a stronger leak (see (1) and starter neuron in the SNN). Consecutive to a spike emission, an electrical flux is sent, transformed at the synapse into a local excitatory or inhibitory synaptic postpotential current. This is then received at the targeted elements (2). The synapse is computationally modeled as a dynamical weight and is subject to be modulated from learning functions. The learning rule we used in this SNN is an adapted spike-timing dependent plasticity (STDP) [30–32]. The result of a STDP function is to increase a synaptic weight if the preneuron spikes before the postneuron unit, in a defined short time window. If the prespike arises after the postspike, then the inverse correlation leads to decreasing the synaptic weight (3).

*Equation 1: Discrete-Time Neural Input Integration Function.* Consider the following:

$$v_m(k) = f(v_m(k-1) + v_i(k)) + l. \quad (1)$$



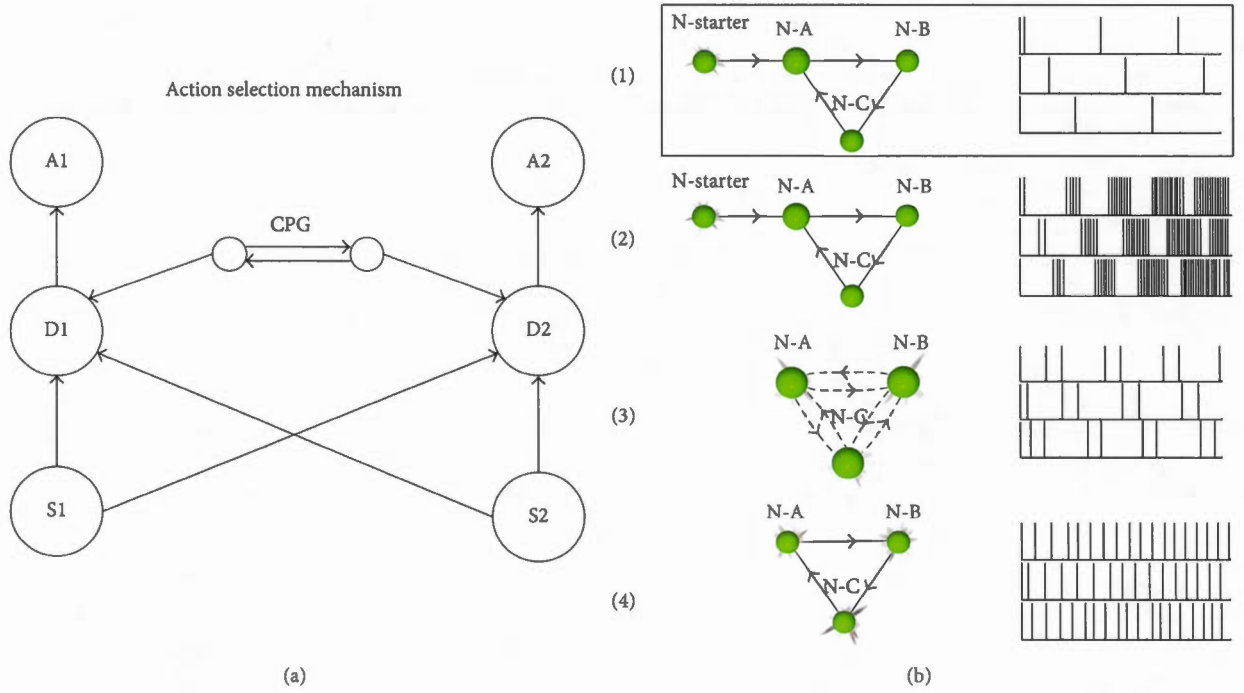


FIGURE 1: Dynamic of the AS mechanism and CPG topologies. (a) First, a stimulus is perceived from a sensory neuron. Then, the spikes emission from the sensory neuron targets all decision neurons, conserving a neutral aspect in the choice of possible actions to come. Finally, when the sensory input coincides with the CPG input at the decision neuron, the linked action will be performed (alternatively switching between actions). (b) In (1), the CPG represents the serial circular excitatory topology we used in this work. In (2), we show one option among several possibilities of tuning differently the internal neural parameters such as the postpotential state value, the refractory period, and the length of the neural curves. In (3), we show an all-to-all reciprocal inhibitory CPG circuit built with endogenous pacemaker neurons. In (4), we still keep the pacemaker neurons in the CPG but it is now linked with excitatory synapses. Overall, (b) (comparative spike bars in the same time scale on the right side) demonstrates the diversity of the output patterns reflected in different CPG, derived from small structural and functional changes.

$v_m(k)$  = membrane potential at cycle  $k$ ,  $v_t(k)$  = sum of the synaptic input as calculated in (2),  $f$  = ascending exponential function set between 0 and threshold (set as 65), and  $l$  = leak current for pacemaker property (set as 1).

**Equation 2: General Alpha Function Representing the Postsynaptic Potential Curve.** Consider

$$f(t) = ge^{-t/\pi}. \quad (2)$$

$g$  = amplitude (set as 20),  $\pi$  = tau (set as 7), and  $t$  = time since spike (in cycle).

**Equation 3: STDP Function Used.** Consider the following:

$$\Delta w = \alpha_{t_{\text{post}} - t_{\text{pre}}} e^{(t_{\text{post}} - t_{\text{pre}})/\pi}. \quad (3)$$

$\Delta w$  = synaptic weight change,  $\alpha_{t_{\text{post}} - t_{\text{pre}}} = 1$  or  $-1$ , depending on the sign of  $t_{\text{post}} - t_{\text{pre}}$ , and  $\pi$  = time constant.

STDP coefficient for  $\Delta w$  is as follows.

Maximum variation period = 3000 cycles.

Maximum synaptic change = 35%.

Maximum STDP time window = 25 cycles.

**2.2. The AS Process.** The elements in the AS mechanism consist of four basic groups of cells. The first group represents the decisions or command-neurons pointing to the action neurons which activates actuators (second group). The third group contains the sensory neurons providing contextual inputs linked to the decision neurons. Finally, the last group of cells contains the CPG neurons weakly connected to the decision neurons. The main function of the CPG is to provide a regular oscillation output pattern to bias one preferred decision neuron over the others. Notice that, a CPG neuron output is never allowed to trigger its linked decision neuron since the EPSP is too weak to reach the spike threshold. However, when pairing sensory and CPG inputs, only then can it reach its threshold and spike (see Figure 2). Therefore, the tuning of the parameters must overlap in relation of the CPG period and the sensory duration. In our experiments, a full CPG loop takes 90 cycles; hence, one CPG neuron spikes every 30 cycles. The sensory input duration last approximately 110 cycles. A second effect of the CPG is to disambiguate equal sensory inputs, a known conflicting problem difficult to resolve in the AI domain. Finally, CPG could also be understood as rhythmic internal values, feeding input in the AS process.

To graphically represent the AS process (Figure 1, left side), we show it in a complete generic scenario of two sensors

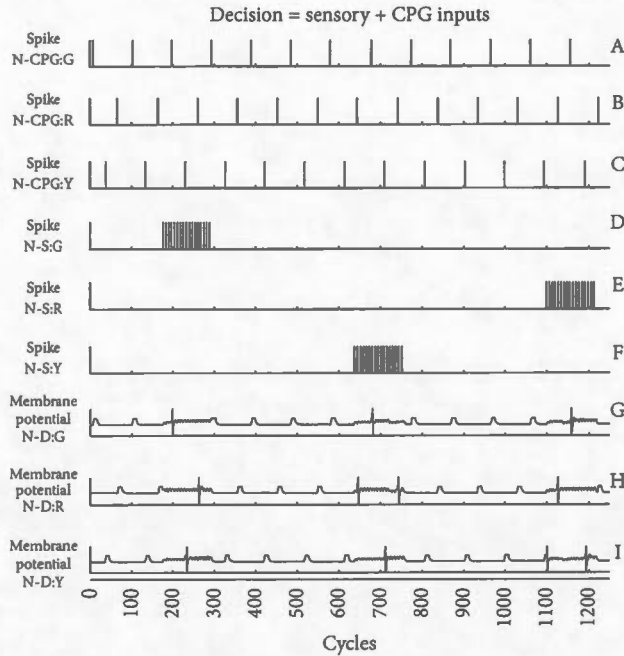


FIGURE 2: The figure shows the regular output spikes pattern from the serial circular excitatory CPG neurons (graphics A to C), which is the exact topology we used in our experiments. The mathematical neural dynamic behind these neurons is the same as the other spiking cells. When the CPG spike coincides with the sensory spikes (graphics D to F), the membrane potential of the decision neurons (graphics G to I) reaches the threshold for a spike emission (see the AS process in a complete SNN architecture 4).

and two actions. The SNN architecture is divided into three distinct layers: the sensory inputs, the internal integrative states, and the external action outputs. In Figure 3, the AS components are also clustered in a singular module within the generic but detailed SNN.

For the CPG's kernel, we chose to embed the most regular and minimalist structure (see option 1 in the highlighted right side of Figure 1). The synaptic weights were all set to 100%, in order to have a continuous spike loop. To start the CPG, we used a biologically plausible endogenous pacemaker that shuts down just after initiating the dynamic. This starter option could be understood in terms of an internal value (i.e., low batteries, attentional process, and sensory-motor input) or could also be any other kind of triggers. As a result, the three neurons of the CPG are stimulated one after the other because of the circular serial excitatory connections 2.

**2.3. The OC Learning Procedure.** The cellular components included in the OC process consist in sensor neurons that provide the contextual inputs for Decision-to-Action neurons to generate the behaviors. Also, external reinforcer pointing to predictor neurons are also connected to Decision-to-Action neurons. Since sensor neurons are weakly linked to the predictor neurons but contain an STDP rule, the repetitive coincidence of the reinforcer (following the desire action) and the sensory input at the predictor neurons will increase

the synaptic weight. Therefore, sensory inputs will eventually trigger actions without any further needs of reinforcers [29].

**2.4. The SNN Architecture.** Specifically used for our results in a three-sensor and three-action context (Figure 4), the sensory neurons are composed of three color sensors (green, yellow, and red) in addition to one light sensor to perceive the rewarding light. The motor output neurons are represented by three LEDs (green, yellow, and red). Our AS process includes, as a modulating element, CPG placed at the intermediary neural layer. It contains three neurons paired with the same number of the possible actions. The proposed CPG kernel consists in excitatory neurons organized in a serial circular topology.

Each CPG neuron in the CPG network is connected to its own decision neuron with a small synaptic weight. One target of a decision neuron is its connected action neuron with a strong synaptic weight between these units; when a decision neuron spikes, the linked action neuron spikes as well. Each decision neuron is also weakly connected to its own predictor neuron for the learning context interrelation. In this experiment, the predictor neurons target their output to all other decision neurons with inhibitory strong synaptic links. Therefore, when the sequence sensor-action-reward is learned from a precise predictor neuron, it will shut down all other possible actions. This arbitrator mechanism could be understood as a type of neural competition. Initial synaptic values used in our SNN were manually tuned and can be retrieved in Table 1.

**2.5. The Task and the Actions.** In the virtual experiment (Figure 5), the SNN is implemented in a static robot. The robot's task consists in learning to match colors between its three possible actions of LED emission (green, yellow, and red) and the color blocks perception. Our 3D simulation software environment (SIMCOG-NeuroSim, AI-Future) allows three different color blocks (green, yellow, and red) to move continuously, at a constant speed, in a clockwise circular trajectory, passing one at a time just in front of the robot. The time frame perceptive contact enables the robot to, at least, produce one different action for each block over 1000 cycles. In the first part of the experiment (0–10000 cycles), a rewarding light (not shown) triggers only when the color LED action matches the block of the same color. At cycle 10500, the robot was moved temporarily for 3000 cycles (cycle 10500–13500) to a location where no sensory input is received, allowing a forgetting factor to operate and reset the synaptic weights. Then, the robot was replaced to its initial position for another round. However, in this part, the rewarding light follows only when the LED emission is on the next color block. The purpose of this part is to show how efficient the AS and OC dynamics can modify the behavior, since these novel learning associations are achieved in a single trial.

As a proof of concept and endpoint in the robotic domain, we reproduced the virtual setup in a physical experiment (Figure 6). The SNN is totally identical and we simply transferred it into the physical robot without any further adjustments. For simplicity, we chose the EV3 Lego Mindstorm (Lego Inc.) as physical platform. The main processor is an

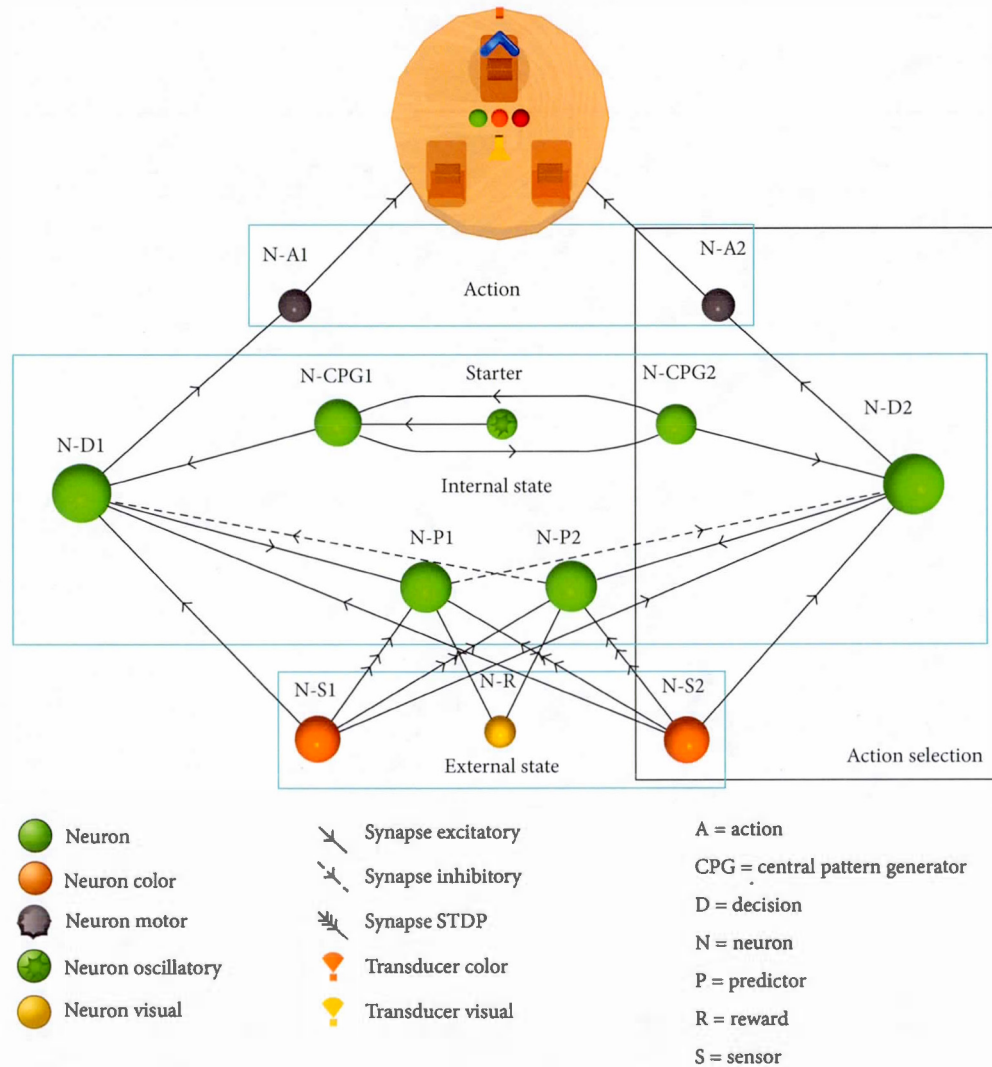


FIGURE 3: Generic SNN of AS and OC in a two-sensor and two-motor configuration.

ARM9 core clocking at 300 MHz and it contains 64 MB RAM. The LEDs are similar to the virtual scenario except that there are only two colors available, the green and the red. When the two of them are opened at the same time, the resulting color is orange, hence having our third color for the experimentation. A light sensor is also used to read the external rewarding light, which was synchronized and delivered from a Raspberry Pi board, just after a desired action is done by the robot. A NXT Lego Mindstorm controller (Lego Inc.), mounted on a shaft, controls the rotation of three color bricks (green, orange, and red) using one attached motor. A slow stepwise speed was set with no possibility of modulation from the robot. In this configuration, the bricks pass just in front of the color sensor. When a sensor catches a color block, the numerical value is converted in an artificial electrical current with an adapted scaling factor for the SNN. Only the first learning part was done for the demonstration. Supplementary material is available at <https://www.youtube.com/watch?v=8MXA4wxJSPE> and consists of a video of the experiment.

### 3. Results

The results from the virtual experiment were obtained in a single trial. The following graphic data will refer to Figure 4 for the SNN architecture and Table 1 for its associated synaptic weight matrix. In Figure 7, we can observe at the beginning of the simulation that when the Green Sensor neuron (N-S:G) spikes (black bars in graphic A), the robot tries alternative actions of lighting up each LED (graphic B: green, graphic G: yellow, and graphic L: red). Since there was no reward for any actions triggered prior to cycle 300, no learning from the STDP rule was observed at the synapse going from the sensor to the predictor neuron (D, I, and N). At around cycle 500, a first yellow block is perceived from the sensor yellow (F), while the CPG continuously provides alternative actions of LED emissions. Specifically, with the lighting up of the yellow LED (G), and with the following light reward (not shown), the associated predictor neuron spikes (H). Consequently, a positive association between the Yellow



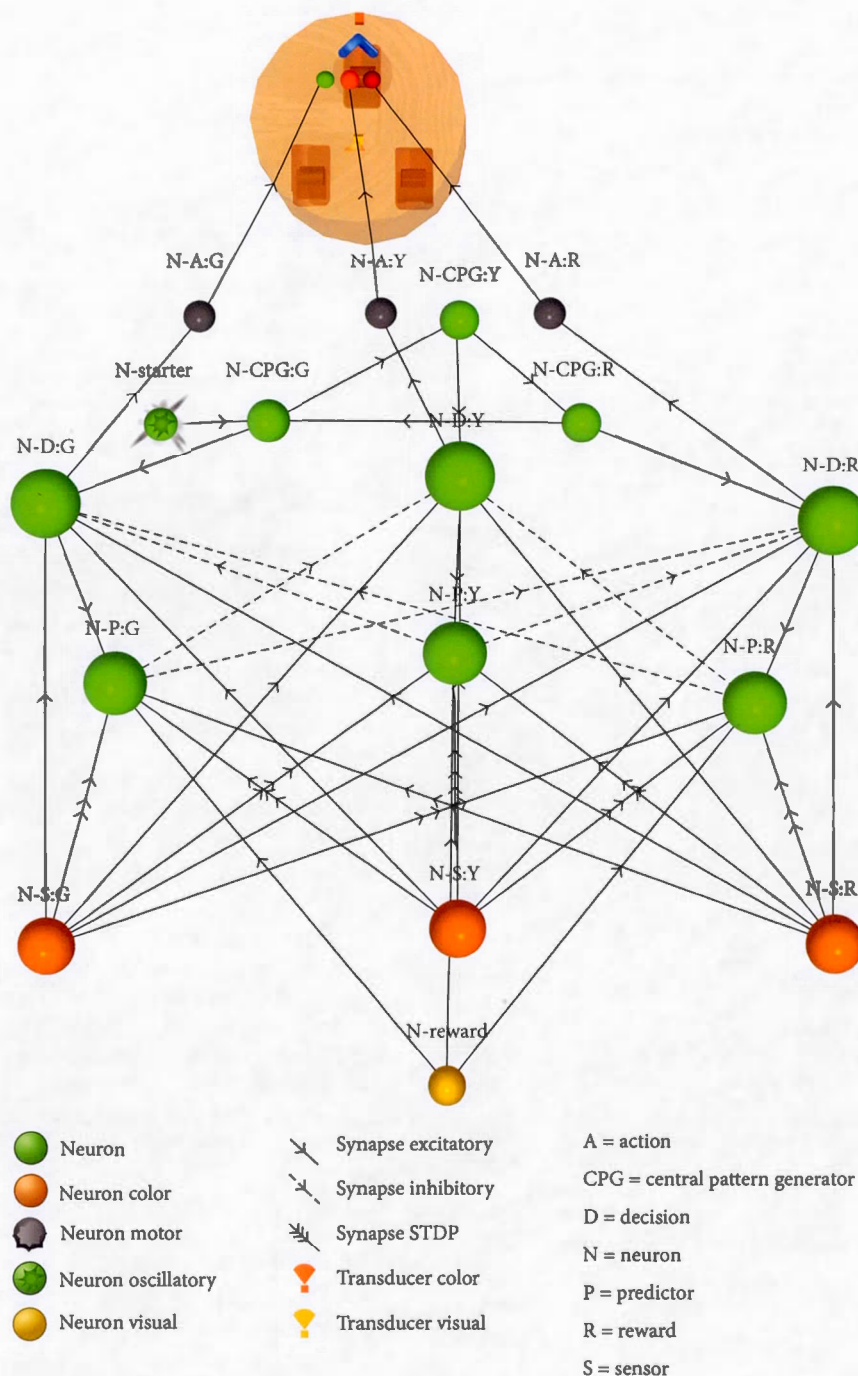


FIGURE 4: SNN architecture used for our experiments, demonstrating the interrelation of the AS process and the OC learning rule in a three-sensor and three-motor configuration. See also Figure 3: A = action, CPG = central pattern generator, D = decision, G = green, N = neuron, P = predictor, S = synapse, R = red, Y = yellow, and STDP = three consecutive arrows on a synapse.

Sensor neuron and this predictor neuron starts to increase the STDP coefficient (I). This affects the synaptic weight to a bound limit when several associations occurred, stabilizing at around cycle 4000. The role of the predictor neuron (H) in this SNN is to inhibit the other decision neurons and their connected action neurons (B, L). At around cycle 8500, one

can see that the robot has fully learned the three sensory-motor contexts by pairing the good LED action with the good perceived color block. Since the period of the CPG neurons and the rotation of the color blocks did not fit perfectly, the learning time frame for each sensory-motor pair is not identical.



TABLE 1: Synaptic weights associated with the SNN of Figure 4. These are represented by numerical percentage values. Delays in cycle as well as the type of the synapse are also mentioned.

Synaptic weight table			
Name	Strength	Delay	Type
S-A:G/E:G	5	0	Excitatory
S-A:R/E:R	5	0	Excitatory
S-A:Y/E:Y	5	0	Excitatory
S-CPG:G/CPG:Y	100	30	Excitatory
S-CPG:G/D:G	30	0	Excitatory
S-CPG:R/CPG:G	100	30	Excitatory
S-CPG:R/D:R	30	0	Excitatory
S-CPG:Y/CPG:R	100	30	Excitatory
S-CPG:Y/D:Y	30	0	Excitatory
S-D:G/A:G	100	0	Excitatory
S-D:G/P:G	15	0	Excitatory
S-D:R/A:R	100	0	Excitatory
S-D:R/P:R	15	0	Excitatory
S-D:Y/A:Y	100	0	Excitatory
S-D:Y/P:Y	15	0	Excitatory
S-Reward/P:G	8	0	Excitatory
S-Reward/P:R	8	0	Excitatory
S-Reward/P:Y	8	0	Excitatory
S-S:G/D:G	15	0	Excitatory
S-S:G/D:R	15	0	Excitatory
S-S:G/D:Y	15	0	Excitatory
S-S:G/P:G	15	0	Excitatory
S-S:G/P:R	15	0	Excitatory
S-S:G/P:Y	15	0	Excitatory
S-S:R/D:G	15	0	Excitatory
S-S:R/D:R	15	0	Excitatory
S-S:R/D:Y	15	0	Excitatory
S-S:R/P:G	15	0	Excitatory
S-S:R/P:R	15	0	Excitatory
S-S:R/P:Y	15	0	Excitatory
S-S:Y/D:G	15	0	Excitatory
S-S:Y/D:R	15	0	Excitatory
S-S:Y/D:Y	15	0	Excitatory
S-S:Y/P:G	15	0	Excitatory
S-S:Y/P:R	15	0	Excitatory
S-S:Y/P:Y	15	0	Excitatory
S-Starter/CPG:G	50	0	Excitatory
S-P:G/D:R	10	0	Inhibitory
S-P:G/D:Y	10	0	Inhibitory
S-P:R/D:G	10	0	Inhibitory
S-P:R/D:Y	10	0	Inhibitory
S-P:Y/D:G	10	0	Inhibitory
S-P:Y/D:R	10	0	Inhibitory

Action selection and operant conditioning

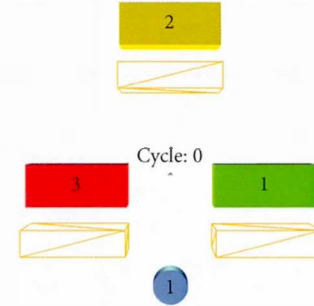


FIGURE 5: Snapshot of the 3D environment including the robot and three color blocks rotating constantly clockwise and passing just in front of the color sensor of the robot.

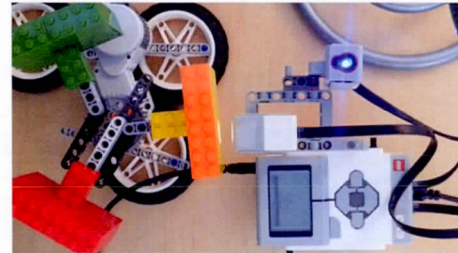


FIGURE 6: Picture of the EV3 brick Lego Mindstorm and the rotative color blocks setup. The color sensor points to the blocks and the light sensor points up.

Between cycle 10500 and 13500, we changed the robot's location to avoid perception of the color blocks. This was done in order to allow the SNN to reset the synaptic weights to their initial values (using a forgetting parameter present in the STDP rule). This was optional, and learning forever would be the result scenario if the feature was not active. Unlearning could also be obtained from inverting the temporal sequence of the sensor, action, and reward. If there is no correlation anymore, the STDP rule will progressively decrease the synaptic weight. In another simulation setup, a punishment (inhibition) could also serve as a fast negative modulation factor of the synaptic weights.

The last part (>13500 cycle) of Figure 7 demonstrates the online adaptive behavior aspect of the SNN embedding OC and AS. One can observe that the robot must choose a different action in order to receive the reward. In this case, lighting up a green LED on a yellow block, a yellow LED on a red block, and a red LED on a green block triggers the reward. The corresponding STDP factors (E, J, and N) match these three learning sets.

As for our physical simulation, the architecture was not modified in any way, except for the binding of logical sensors and motors to the robot. Figure 8 shows the results of the simulation, which were obtained in a single trial. They show approximatively the same data, with more or less precision and small artifacts. This is due to the fact that it is indeed much easier to configure variables and the context of virtual environments than it is in the real world.

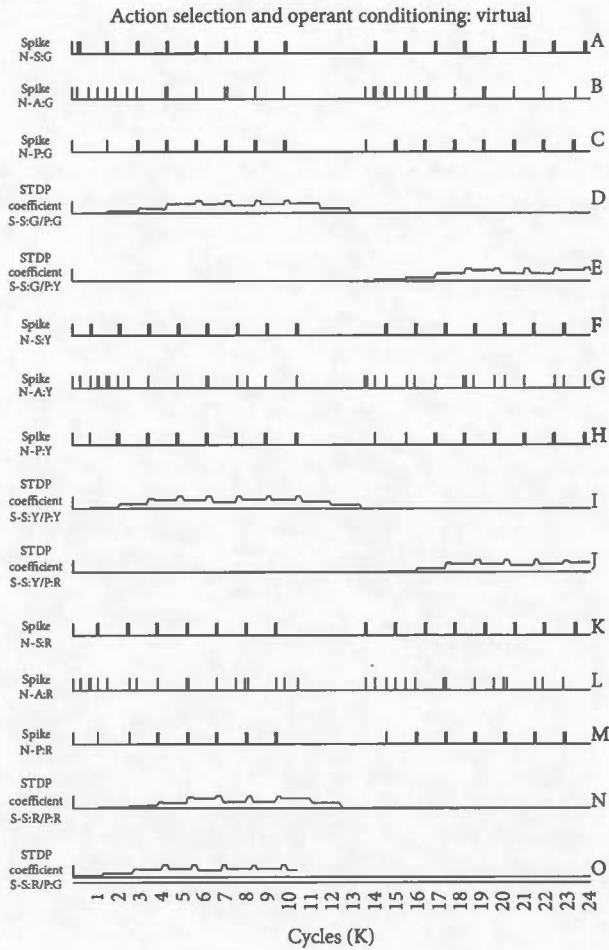


FIGURE 7: Graphics of the neural spikes and the STDP rule factors occurring over 24000 cycles. The first five (graphics A to E) concern the green block and green LED. The middle five capture the logic of the yellow block and the yellow LED. The last group of five is associated with the red color. Around cycle 8500, the robot has learned to perform the right action with the appropriate cue. This means that, upon triggering an action, if the predictor spikes from a following reward, the STDP coefficient will increase, boosting the synaptic weight between the sensor and the predictor. Once this synaptic weight reaches a critical threshold, the sensor input will trigger a spike to the reinforced predictor neuron, without the need of the reward.

#### 4. Discussion

In this paper, we explored the AS process through a neuro-robotic perspective. Since this general mechanism directly involves actions, we demonstrated the phenomenon in the context of OC procedures which also imply a selection of actions from reinforcer. Our main objective was to study the benefit effects of merging this learning rule with an AS process. A second concern was to provide a fast track solution to efficiently design more complex SNN used as brain's controller for virtual and physical robots that include several motor outputs. We propose a basic CPG motif as one

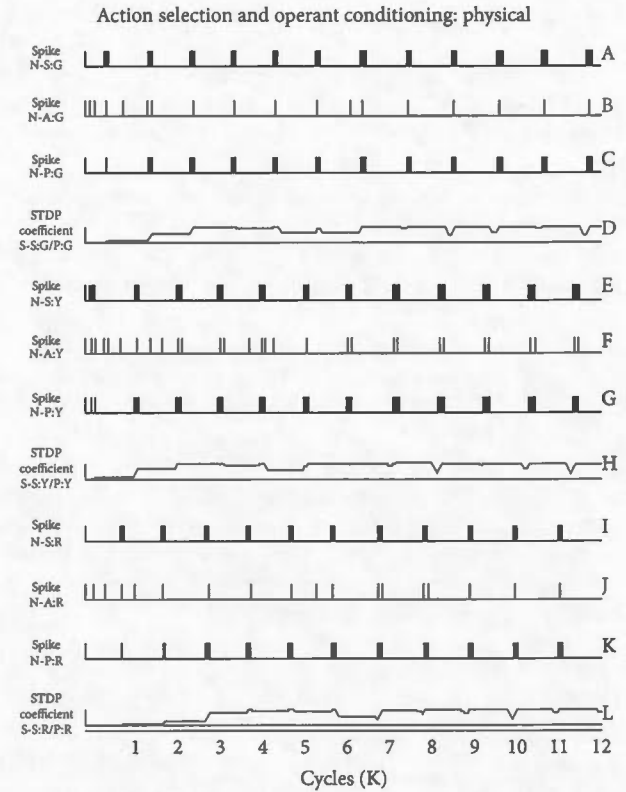


FIGURE 8: Graphics of the neural spikes and the STDP rule factors occurring over 12000 cycles.

key component of an AS process, in order to neutrally switch between its available actions. With the CPG structure used in relation with a sensory input context, a decision neuron gets all the information needed to bias toward one preferred action. We also showed that the OC learning function influences the AS process, conferring supplementary adaptive behaviors from synaptic plasticity.

We chose a simple CPG topology as one component of the AS mechanism. Other CPG configurations are possible [33], including those built with reciprocal inhibitory synaptic links and endogenous pacemaker neurons, though the analytical issues are more complex to track and predict. Tuning the parameters (i.e., postpotential spike value, threshold) of individual neuron differently could also influence the rhythm, affecting the CPG network by increasing or decreasing their output periods. After several options, we found that a serial excitatory circular CPG motif is a good trade-off between simplicity and benefits. In our AS model, without any other synaptic feed, this CPG configuration will spike one unit after the other, indefinitely and at a constant rate. We showed the AS model in a generic two-two example and in a specific three-three sensory input and motor output configurations. Adding more sensors and actions will necessary require other neurons in the CPG network, though their numbers are linearly related to their attached decision and action neurons, acting as a premotor structures. In this case, hierarchical

groups of CPG/actions could also replace the serial circular topology, possibly avoiding useless spikes or triggering other networks. Allowing different combinations and compositions of CPG units also dramatically increase the behavior possibilities, without considering a one-to-one CPG-action, though it was not explored in this paper.

Adding the decision neurons (equivalent to command-neuron in invertebrates) into the AS model allowed flexibility, regarding several contextual input sources. The CPG units bring the decision neurons membrane potential to a subthreshold firing level. Since the CPG period is fixed, the speed-accuracy trade-off (SAT) of the decision-making [34] result is fast and accurate. Unfortunately, but no adjustment is possible, a major point to consider when modeling an AS process. Sometimes, cognitive agents must take decisions quickly while in other conditions, it is necessary to take the time to compute the best decision. According to a recent hypothesis [35], flexibility of the SAT's response variables depends on adjusting the baseline firing rate, the sensory gain and noise inputs, the firing threshold, and related bound parameters of the receiving neurons [25, 36].

Having those various AS modulator factors in mind, designing complex SNN with several populations of neurons including heterogeneous individual neural parameters values is possible, but highly complex to tune properly. In this perspective, progressively integrating stronger/lower and faster/slower CPG inputs could add discriminative and flexible response advantages as well as offering more realistic behavioral features of the AS mechanism. A computational challenge in an AS model within the neurorobotic field is to allow the SNN to dynamically change all these initial fixed parameters values, conferring considerable adaptive properties at the level of the cognitive agent. In this vein, perhaps a question remains about the SAT: what are the variables biases in the AS process when there is no emergency to choose one action?

In our experiments, reversing the rotation order for the block to counter clockwise, accelerating or decelerating the speed of rotation as well as mixing the color order do not change the qualitative aspect of the learning curve. However, the temporal relation between the perception time contact of the sensor inputs and the timing of the CPG influences the number of occurrences of these associations and, thus, the length of time needed to learn. In any case, the EPSP timing between the CPG and the sensory neurons is of major importance and needs a full coherence into the whole dynamic system. In this perspective, the physical experiment shed some light on the temporal robustness of our AS and OC models, justifying its inclusion in the study. Without changing any parameters in the SNN, the EV3 robot was able to learn very well how to receive the reward when achieving the good action, even if the rotation of the blocks was irregular due to the imprecision of the material.

We explained how the OC learning rule modulates the AS process in a SNN paradigm. The sensory-motor context does influence the decision to do one action over others. These decisions were not just built-in reflexes. Moreover, the behavioral plasticity was observed even if the CPG dynamic was fixed. At this point, some interesting variances could be to add

other learning rules that are not limited by an OC procedure. As such, integrating nonassociative (habituation) and other associative learning (classical conditioning) functions could complete the design of an AS model, but it was beyond the scope of this paper.

One concern we avoid in this paper is the attentional problem. We chose to ignore this major cognitive component, mostly because of the current lack of neural mechanism and theory when applied in lower neural system organisms. We understand that basal ganglion or any subcortical or cortical implications in the AS process seem to be relevant structures in higher biological neural systems such as in humans or primates. However, our present perspective on the AS problem lies in the AI neurorobotic domain, which is still beyond the reach of lower cognitive natural species. Our strategy aims to engineer a bioinspired minimalist solution from emulating simple neural organisms such as in *C. elegans*, which selects its actions without involving huge structures. Instead, command-neurons and CPG neurons are common cellular elements found in the primitive invertebrate neural circuits. No doubt, complex neural layers may extend and add profitable values in the simulation of an AS process but should not be a necessary requirement to achieve a basic one. These evolutionary concerns may eventually find an echo in a multiplicity and hierarchical AS mechanisms.

The AS scope obtained from these results is theoretically not limited to only a few simple actions or unimodal single sensory stimuli. The generic aspect of the AS process comes from the parsimonious components and parameters inside the kernel. The simplicity of this AS module already allows to be adaptive from a  $2 \times 2$  to a  $3 \times 3$  sensors-actions scenario without much changes in the SNN architecture. In these two scenarios, as long as there are the same numbers of possible ending actions and sensors as input, the AS core process will operate and tune the same. Therefore, building more complex SNN including several actions should be anticipated as faster and easier, though it remains to be proven in other situations. In that sense, we currently work on a shaping behavior learning technique based on the AS process, while simulating an indoor dynamical navigation task with several possible behaviors. This is an example that can demonstrate how this bioinspired AS process could help in concrete application in the robotic field.

## 5. Conclusion

This paper showed an AS process made from simple cellular elements. It is based on CPG and sensory neurons which influence decision neurons in their choice to generate a behavior from the action neurons. We demonstrated this basic AS mechanism in an OC learning context that allows behavioral flexibility from their mutual influences. The experiments were conducted under a biologically inspired paradigm, specifically with a SNN acting as brain-controller for virtual and physical robots. In addition, the simplicity and the generic aspect of our AS model may provide a fast track solution to build more complex SNN, including multiple actions in different dynamic scenarios.



## Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

## References

- [1] A. K. Seth, T. J. Prescott, and J. J. Bryson, *Modelling Natural Action Selection*, Cambridge University Press, Cambridge, UK, 2011.
- [2] M. D. Humphries, K. Gurney, and T. J. Prescott, "Is there a brain-stem substrate for action selection?" *Philosophical Transactions of the Royal Society B: Biological Sciences*, vol. 362, no. 1485, pp. 1627–1639, 2007.
- [3] J. I. Gold and M. N. Shadlen, "The neural basis of decision making," *Annual Review of Neuroscience*, vol. 30, pp. 535–574, 2007.
- [4] M. Khamassi, L. Lachèze, B. Girard, A. Berthoz, and A. Guillot, "Actor-critic models of reinforcement learning in the basal ganglia: from natural to artificial rats," *Adaptive Behavior*, vol. 13, no. 2, pp. 131–148, 2005.
- [5] N. F. Lepora, C. W. Fox, M. H. Evans, M. E. Diamond, K. Gurney, and T. J. Prescott, "Optimal decision-making in mammals: insights from a robot study of rodent texture discrimination," *Journal of the Royal Society Interface*, vol. 9, no. 72, pp. 1517–1528, 2012.
- [6] T. J. Prescott, F. M. M. González, K. Gurney, M. D. Humphries, and P. Redgrave, "A robot model of the basal ganglia: behavior and intrinsic processing," *Neural Networks*, vol. 19, no. 1, pp. 31–61, 2006.
- [7] W. Maass, "Networks of spiking neurons: the third generation of neural network models," *Neural Networks*, vol. 10, no. 9, pp. 1659–1671, 1997.
- [8] T. C. Stewart, T. Bekolay, and C. Eliasmith, "Learning to select actions with spiking neurons in the basal ganglia," *Frontiers in Neuroscience*, vol. 6, article 2, 2012.
- [9] S. Skorheim, P. Lonjers, and M. Bazhenov, "A spiking network model of decision making employing rewarded STDP," *PLoS ONE*, vol. 9, no. 3, Article ID e90821, 2014.
- [10] E. M. Izhikevich, "Simple model of spiking neurons," *IEEE Transactions on Neural Networks*, vol. 14, no. 6, pp. 1569–1572, 2003.
- [11] S. Faumont, T. H. Lindsay, and S. R. Lockery, "Neuronal micro-circuits for decision making in *C. elegans*," *Current Opinion in Neurobiology*, vol. 22, no. 4, pp. 580–591, 2012.
- [12] M. Zhang, W. R. Schafer, and R. Breitling, "A circuit model of the temporal pattern generator of *Caenorhabditis* egg-laying behavior," *BMC Systems Biology*, vol. 4, no. 1, article 81, 2010.
- [13] A. L. Stöckl, R. Petie, and D.-E. Nilsson, "Setting the pace: new insights into central pattern generator interactions in box jellyfish swimming," *PLoS ONE*, vol. 6, no. 11, Article ID e27201, 2011.
- [14] C. R. von Reyn, P. Breads, M. Y. Peek et al., "A spike-timing mechanism for action selection," *Nature Neuroscience*, 2014.
- [15] S. Grillner, "Neurobiological bases of rhythmic motor acts in vertebrates," *Science*, vol. 228, no. 4696, pp. 143–149, 1985.
- [16] K. Matsuoka, "Mechanisms of frequency and pattern control in the neural rhythm generators," *Biological Cybernetics*, vol. 56, no. 5–6, pp. 345–353, 1987.
- [17] A. J. Ijspeert, "Central pattern generators for locomotion control in animals and robots: a review," *Neural Networks*, vol. 21, no. 4, pp. 642–653, 2008.
- [18] M. Heisenberg, "Action selection," in *Invertebrate Learning and Memory*, R. Menzel and P. Benjamin, Eds., vol. 22, Academic Press, 2013.
- [19] P. Redgrave, T. J. Prescott, and K. Gurney, "The basal ganglia: a vertebrate solution to the selection problem?" *Neuroscience*, vol. 89, no. 4, pp. 1009–1023, 1999.
- [20] P. Cisek, "Making decisions through a distributed consensus," *Current opinion in neurobiology*, vol. 22, no. 6, pp. 927–936, 2012.
- [21] P. Cisek, G. A. Puskas, and S. El-Murr, "Decisions in changing conditions: the urgency-gating model," *The Journal of Neuroscience*, vol. 29, no. 37, pp. 11560–11571, 2009.
- [22] S. D. Brown and A. Heathcote, "The simplest complete model of choice response time: linear ballistic accumulation," *Cognitive Psychology*, vol. 57, no. 3, pp. 153–178, 2008.
- [23] M. Usher and J. L. McClelland, "The time course of perceptual choice: the leaky, competing accumulator model," *Psychological Review*, vol. 108, no. 3, pp. 550–592, 2001.
- [24] M. Kassim, N. Zainal, and M. Arshad, "Central pattern generator in bio-inspired simulation using matlab," in *Proceedings of the MEDINFO*, vol. 98, 1998.
- [25] D. Standage, D. Wang, and G. Blohm, "Neural dynamics implement a flexible decision bound with a fixed firing rate for choice: a model-based hypothesis," *Frontiers in Neuroscience*, vol. 8, article 318, 2014.
- [26] B. Brembs, F. D. Lorenzetti, F. D. Reyes, D. A. Baxter, and J. H. Byrne, "Operant reward learning in *Aplysia*: neuronal correlates and mechanisms," *Science*, vol. 296, no. 5573, pp. 1706–1709, 2002.
- [27] K. M. Crisp and K. A. Mesce, "Beyond the central pattern generator: amine modulation of decision-making neural pathways descending from the brain of the medicinal leech," *Journal of Experimental Biology*, vol. 209, no. 9, pp. 1746–1756, 2006.
- [28] T. M. Wright Jr. and R. L. Calabrese, "Patterns of presynaptic activity and synaptic strength interact to produce motor output," *The Journal of Neuroscience*, vol. 31, no. 48, pp. 17555–17571, 2011.
- [29] A. Cyr, M. Boukadoum, and F. Thériault, "Operant conditioning: a minimal components requirement in artificial spiking neurons designed for bio-inspired robot's controller," *Frontiers in Neurobotics*, vol. 8, article 21, 2014.
- [30] N. Caporale and Y. Dan, "Spike timing-dependent plasticity: a Hebbian learning rule," *Annual Review of Neuroscience*, vol. 31, pp. 25–46, 2008.
- [31] A. Cyr and M. Boukadoum, "Classical conditioning in different temporal constraints: an STDP learning rule for robots controlled by spiking neural networks," *Adaptive Behavior*, vol. 20, no. 4, pp. 257–272, 2012.
- [32] H. Markram, W. Gerstner, and P. J. Sjöström, "A history of spike-timing-dependent plasticity," *Frontiers in Synaptic Neuroscience*, vol. 3, article 4, 2011.
- [33] A. I. Selverston, "Invertebrate central pattern generator circuits," *Philosophical Transactions of the Royal Society B: Biological Sciences*, vol. 365, no. 1551, pp. 2329–2345, 2010.
- [34] P. Cassey, A. Heathcote, S. D. Brown, and O. Sporns, "Brain and behavior in decision-making," *PLoS Computational Biology*, vol. 10, no. 7, Article ID e1003700, 2014.
- [35] R. Bogacz, E.-J. Wagenmakers, B. U. Forstmann, and S. Nieuwenhuis, "The neural basis of the speed-accuracy tradeoff," *Trends in Neurosciences*, vol. 33, no. 1, pp. 10–16, 2010.
- [36] R. P. Heitz and J. D. Schall, "Neural mechanisms of speed-accuracy tradeoff," *Neuron*, vol. 76, no. 3, pp. 616–628, 2012.





## ANNEXE F

### ARTICLE 5 : SIMCOG

#### NEURAL COMPUTATIONS & APPLICATIONS

# AI-SIMCOG: a simulator for spiking neurons and multiple animats' behaviours

André Cyr · Mounir Boukadoum · Pierre Poirier

Received: 20 July 2008 / Accepted: 27 February 2009 / Published online: 17 March 2009  
© Springer-Verlag London Limited 2009

**Abstract** Designing a biologically inspired neural architecture as a controller for a complete animat or physical robot environment, to test the hypotheses on intelligence or cognition is non-trivial, particularly, if the controller is a network of spiking neurons. As a result, simulators that integrate spike coding and artificial or real-world platforms are scarce. In this paper, we present artificial intelligence simulator of cognition, a software simulator designed to explore the computational power of pulsed coding at the level of small cognitive systems. Our focus is on convivial graphical user interface, real-time operation and multilevel Hebbian synaptic adaptation, accomplished through a set of non-linear dynamic weights and on-line, life-long modulation. Inclusions of transducer and hormone components, intrinsic oscillator and several learning functions in a discrete spiking neural algorithm are distinctive features of the software. Additional features are the easy link between the production of specific neural architectures and an artificial 2D-world simulator, where one or more animats implement an input–output transfer function in real-time, as do robots in the real world. As a result, the simulator code is exportable to a robot's microprocessor. This realistic neural model is thus amenable to investigate several time related cognitive problems.

**Keywords** Simulator · Artificial intelligence · Spiking neurons model · Learning and memory · Dynamical synapse · Robot

## 1 Introduction

One reason to use simulators in artificial intelligence is to emulate natural processes. Although current simulators only present a small fragment of what real-world entities are capable of, they provide enough clues for researchers to upgrade their knowledge about the general intelligence phenomenon [1] and they are making the possible realisation of virtual [2] and physical artificial life [3].

One research direction in this context is to simulate complete adaptive systems, including artificial entities that can capture relevant input information, integrate it and act upon it. We refer here to the brain, body and environment paradigm [4], where the concept of embodied cognition takes a large place [5–7] and where the central hypothesis is that the intelligence emerges from an equal interaction between the three actors. From this perspective, the observation of animat behaviour in various environmental settings is a requirement for studying intelligence. Hence, software applications that target this specific view of intelligence must focus on top–down cognition interests and may even ignore biological details for the sake of computational efficiency and model tractability.

Before the embodiment trend, attention was solely dedicated to the brain part of an intelligent system, or the controller(s) of an artificial life form. Specifically, the scientific community largely explored artificial neural network (ANN) to better understand biological functions and cognition. The formal artificial neuron model [8], the functional integrate-and-fire spiking neuron model [9] and

---

A. Cyr (✉) · M. Boukadoum  
Computer Science Department,  
Université du Québec à Montréal, Montreal, QC, Canada  
e-mail: acyr@aifuture.com; andre.cyr1@videotron.ca

M. Boukadoum  
e-mail: boukadoum.mounir@uqam.ca

P. Poirier  
Philosophy Department, Université du Québec à Montréal,  
Montreal, QC, Canada  
e-mail: poirier.pierre@uqam.ca



the associative learning rule between neurons [10] had a great impact on the development of early connectionist thinking. Since then, the various architectures of parallel artificial neurons with different learning rules and activation functions that have been developed offer a rich array of models and toolboxes to help investigate intelligence.

Recently, a new generation of ANN, the spiking neural model [11–13], or spiking neural network (SNN), was introduced on the idea of using pulsed coding instead of the traditional rate coding [14]. The underlying assumption is that crucial information is conveyed by the temporal behaviour of individual spikes and that it must be accounted for in neuron models that aim to study dynamical cognitive processes and real-time physical phenomena. Time scales at the millisecond resolution level, synchrony and phase locking are the key parameters when studying neural behaviour at the spike level. Several models and learning rules with computational efficiency proofs have been developed to exploit the new paradigm [15–17].

On the other hand, the SNN literature is not rich with software products that implement rules extended in the temporal domain and reflect specific biological functions in real-time, or online life-long learning such as habituation, sensitisation and short/long-term memory [18–22]. There exist many general-purpose neural simulators to facilitate our understanding of single neuron firing, brain, cognition or intelligence [23–32]. For example, MATLAB [26] does well to model the precise timing of SNNs. Most of the important parameters are available, editable and standard visualisation of results is possible. MATLAB offers a good means to investigate hypotheses on neural models with temporal dimensions. It also represents a good compromise between low-level and more biophysically realistic software toolboxes such as GENESIS [25], NEURON [28] or SNNAP [30], which are more relevant to neurosciences with their ability to investigate neural function from sub-cellular level to small networks and traditional ANN simulators such as SNNs [31]. A common criticism that can be made against all these simulators concerns their non-intuitive GUIs and the difficulty to easily create asymmetric architectures and heterogeneous neural functions, particularly at the level of individual learning rules, synapses and neurons. In addition, they often do not operate in real-time, an important factor for spiking neural models that intend to represent the real world, and one that is necessary for exporting a spiking neural architecture to a robotic platform.

Finally, testing hypotheses in cognition and robotics, especially, potential links between temporality and intelligence, at the level of a complete entity behaving in a dynamical real-world system, requires adapted simulators to investigate this phenomenon. We are not aware of software that are dedicated to the temporal features of SNN

to study intelligence at the global level description of an animat's adaptive behaviour, even if some efforts are in that direction [24, 29]. There is also Webots5 [33], a simulator where emphasis is put on the building side of a robot, a 3D-world simulator engine and code transfer to some physical robot platforms. It accepts various kinds of predefined controller architecture but, unfortunately, it lacks the GUI to build one, particularly, for SNN systems.

In this paper, we describe a novel software simulator that specifically addresses the temporal issues mentioned above. In particular, it has a user friendly GUI that enables the set up of arbitrary SNN architectures within minutes. In addition, a virtual world editor and simulator facilitate the observation of animat behaviours, with a SNN as the central controller, one that implements interesting biological functions in several temporal dimensions. The software is also designed to export all the data produced from simulation and offers some graphical possibilities. In particular, a final executable algorithm is produced as C code, the real outcome of our SNN simulator. The balance of this paper provides a description of the software, a short demonstration application, followed by a discussion and conclusion.

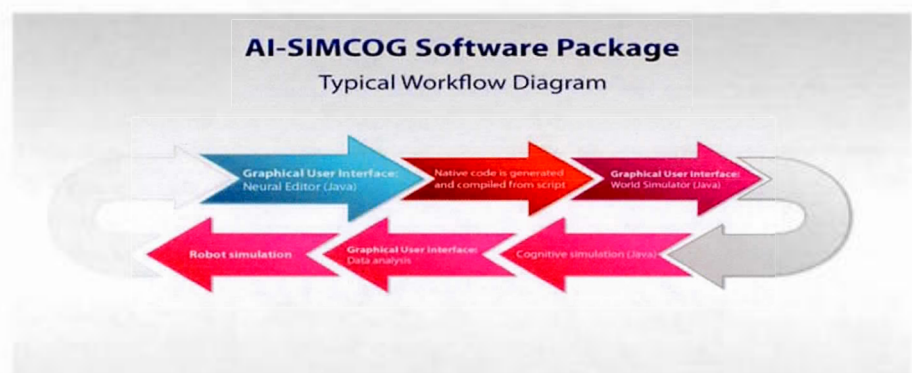
## 2 Description of AI-SimCog

Artificial intelligence simulator of cognition (AI-SIMCOG) is a SNN simulator used to easily design complex real-time SNN architectures, test them on complete autonomous entities in a virtual world and export the corresponding code to a physical robot. It can create and investigate the behaviour of artificial life forms from the level of adaptive dynamical synapses to complex behaviours of one to many complete animats. Special emphasis is made on temporality aspects and robotic implementation issues. The software includes two major packages: a SNN editor and a virtual world simulator-editor. Figure 1 provides the general workflow diagram of the software and Fig. 2 provides the global sequence diagram of a simulation run. Intuitive graphical menus and real-time emulation of SNN activity characterise the application. Extended temporal functions and simulation of biological neural components also represent distinctive features of the software. A description of each component is provided in the next sections.

### 2.1 Neural architecture: GUI editor

The creation of a neural architecture starts by opening a blank project from the main menu of the editor (Fig. 3), assigning physical dimensions to the workspace. In its current implementation, AI-SIMCOG is meant to model small neural architecture comparable to the size of



**Fig. 1** General workflow diagram

*C. elegans*, a well-known low-level animal model in neurosciences [34] that possesses 302 neurons and about 5,000 chemical synapses. This relatively simple neural network is sufficient for *C. elegans* to express many memory properties, basic learning functions and multiple behaviours. In our design, neural architectures of less than 100 neurons, each with synapses, transducers and effectors, are targeted for computational efficiency and given that *C. elegans* represents the level of cognition we want to reach in natural intelligence, and later transpose to an artificial one.

After creating the project, the user enters information about the effectors in four possible configurations: standard four-wheel drive, holonomic three-wheel drive, two-wheel drive and six-wheel drive. Animat shape can also be set (only a circular scalable size has been implemented in the current version). Then, the user organises the animat's neural architecture with a graphical sub-menu where various icons for editable tools are displayed. Neurons, synapses, transducers—sensors, hormones, electrode device<sup>1</sup>, effector motorised wheels and effector motorised turrets<sup>2</sup> are the main items offered in the menu.

Several functions exist to set individual neural activity modes as normal, burst with various built-in kinetic possibilities, and endogenous oscillation (Fig. 4). The normal mode uses a standard action potential curve in response to a constant stimulus. The burst mode can use an arbitrary discharge pattern of three action potentials for the same stimulus. The options of five and ten spikes with decaying amplitudes are also possible. The intrinsic oscillatory or leak mode accounts for neural activity that may occur with no external stimulus, due solely to the non-linear integration of membrane leakage current, and leading the neuron to fire repetitively [35, 36]. The speed and period of oscillation can be adjusted via a set of internal

parameters such as the absolute refractory period. Under certain circumstances related to time-events, complex rhythms could be realised with the inclusion of a special dynamical inner switch, leading to triggering the normal mode through neural pacemaker behaviour. Other conditions may return the activity back to the normal neural mode. This feature is interesting as it allows the creation of endogenous firing patterns that may not be the result of external stimuli. In addition, if an oscillatory neuron is linked to one or more effectors, it can create autonomous and spontaneous movements in an animat or physical robot. Finally, intrinsic oscillation allows the easy creation of central pattern generators and it may serve to model several small local neural circuits that are useful for designing primary locomotion behaviours [37] and even certain circadian phenomena when combined to hormone dependant behavioural issues.

The settable parameters for individual neurons include name, initial membrane potential, rest membrane potential, after-hyperpolarization membrane potential, threshold, absolute refractory period and several non-linear temporal shapes for the membrane potential (see Fig. 5).

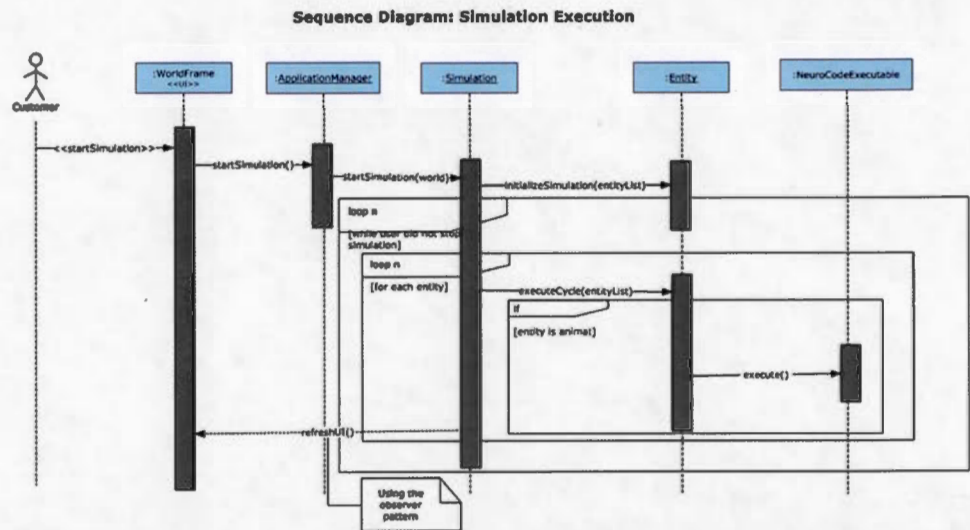
In our simulator, there is no real current to manage, so we do not offer simulating details at the channels or ionic levels. We just simulate current as a steady-state numerical gradient representing the hypothetical difference between inside and outside the artificial neuron. However, as in natural conditions, this gradient is subject to intensity modulation in the time-domain, as required by SNN coding.

The units we use in the simulator are percentages for gradient current and algorithm cycles for time. The execution of one cycle is relative to a hardware real-time clock but can be tuned to average of 1 ms in a standard physical set-up. In fact, running one cycle under the millisecond is a major concern of the software. The highest peak membrane depolarisation arises when there is the emission of a spike, which turns to be 100%, and the maximum of membrane hyperpolarisation is 0%. The neural references for these parameters reflect in some ways the usually accepted

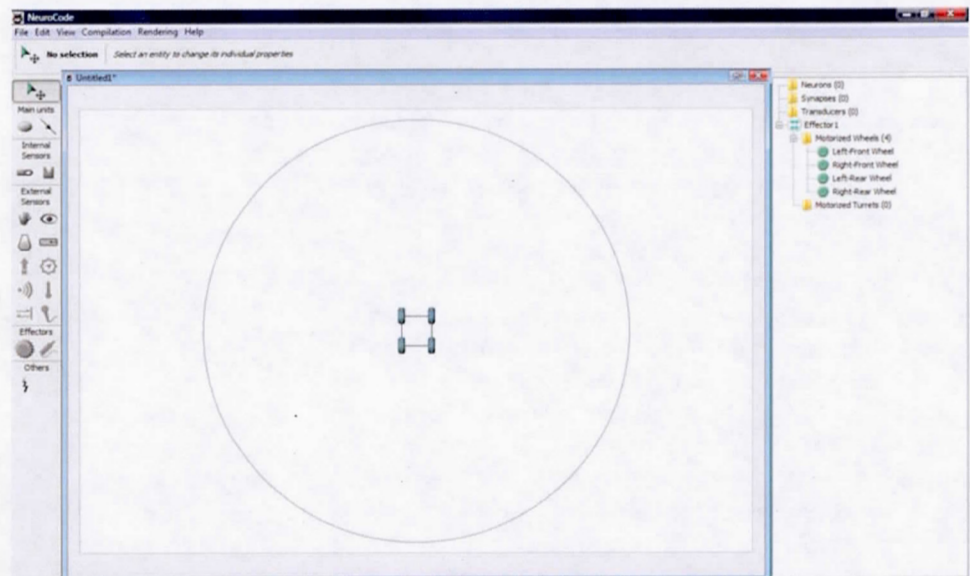
<sup>1</sup> When implemented in a specific neuron, provide a constant positive inward current as a tool to investigate spontaneous neural activity and its contribution to the overall neural dynamics.

<sup>2</sup> Small round structures built as a different layer on top of the animat's external shed, that generate their own independent movements.

**Fig. 2** Sequence diagram of a typical simulation



**Fig. 3** Main window of the neural architecture editor



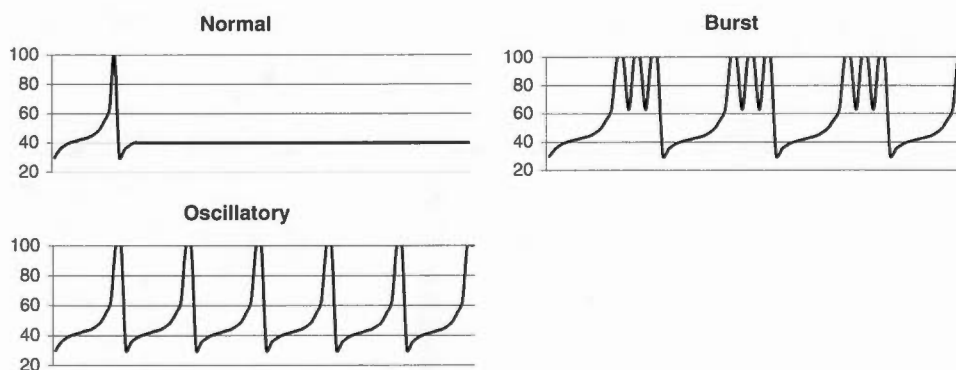
voltage differences that are found in a pyramidal neuron, i.e. a resting membrane potential  $-60$  mV, an after-hyperpolarisation of  $-90$  mV and a depolarisation of  $+40$  mV. The scale is just converted to positive percentage values. The absolute refractory period is counted as a one to three algorithm cycle. The choice of these elementary units allows flexibility for specifying the machine it runs on.

Each parameter is initially set with respect to the standard shape of a formal neuron's kinetic conductance. In usual SNN analytical descriptions, this curve is represented by a set of non-linear equations. During code execution and depending on the accuracy or efficiency sought, some variables are skipped or approximated by linear equations. In AI-SIMCOG, the fluctuations of the membrane potential are initially pre-set and stored in a non-linear scalar vector

whose elements are editable. The software provides three built-in neural kinetic sigmoid curves (Fig 6). The general S-curve is split to around 40% of the amplitude level, with three hyperpolarisation logarithmic curves and three depolarisation exponential curves, creating nine different combinations of neural membrane potential signatures. This feature was introduced to investigate the possible effect of intrinsic neural diversity and heterogeneous neural architecture on behavioural observation.

The functions for individual synapses are split into excitation and inhibition with different strength weights as reflected by the various shapes and variations of the post-synaptic membrane potential, also pre-set in non-linear scalar vector. The variations occur across a time-window that is flexible in intensity (up to 10%) and duration (up to five cycles). Integration of the post-synaptic potential is

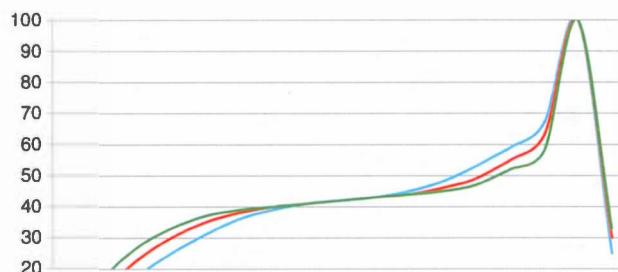


**Fig. 4** Mode of neural membrane potential**Fig. 5** Menu of settable neural parameters

also performed non-linearly through the neural membrane potential and it depends on synapse sign, neural phase and state in the excitation curve and applied learning rules.

One or more synaptic recurrences are allowed on the same neuron, as well as one or more links between neurons (e.g. forward and backward). All parameters are local, including both non-associative learning rules (habituation and sensitisation) and associative learning rules (short and long term memory). These biological functions are not fully understood by the research community, so our intention is only to transpose in software, the essential outcome aspects of the mechanisms. Those are achieved through a set of counter variables tuned and overlapped in different exponential timescale times. Some of them concern association of stimulus and pre-synaptic elements (non-associative memory); others concern pre and post-synaptic elements (associative memory). We are currently investigating the extension of these functions in our software with current knowledge in neurosciences, beyond the simple habituation, short and long-term memory processes. In fact, recent data include several mechanisms for each function like short, intermediary and long habituation behaviours [22] and many facets of long-term potentiation in the specific case of excitatory/inhibitor synapses, normal or burst effect and several asymmetrical window depending on the spike-timing association (Hebb rules) [20].

In addition, three different levels of noise can be added in any combination for each neural activity to mimic the dynamic remodelling factors and the protein turnover of biological synapses. These remodelling parameters or noise functions that we implemented lead to the generation of a transient decrease of the synapses' strength (less than 5%) over many temporal scales. For example, every 50 cycles,

**Fig. 6** Kinetics of neural membrane potential

the synaptic weight connection will decrease by 1% over 2 cycles. Then, every 500 cycles, it will decrease by 2% over 4 cycles and every 10,000 cycles, it will decrease by 3% over 12 cycles. Individual synapses are not modulated at the same time and a random sequence is applied to determine the order of remodelling. All remodelling follows a bell-shaped modulation curve for smooth results. In addition, to avoid a static routine, these modifications may be altered by the dynamics of individual synapses and cell computation activity, which in turn are tributary of the neural architectural setup, the virtual or physical world perception and the actions produced within. This is biologically plausible and offers a good alternative to solely random mathematical equations.

Synapses are the links not only between neurons but also between neurons and effectors, with muscles and neuromuscular junctions representing the natural counterpart. We model this feature with an operator matrix that converts the integration of each post-synaptic potential current to the rotation of motor gears. The translation and effective movements depend on the specific effector-motor

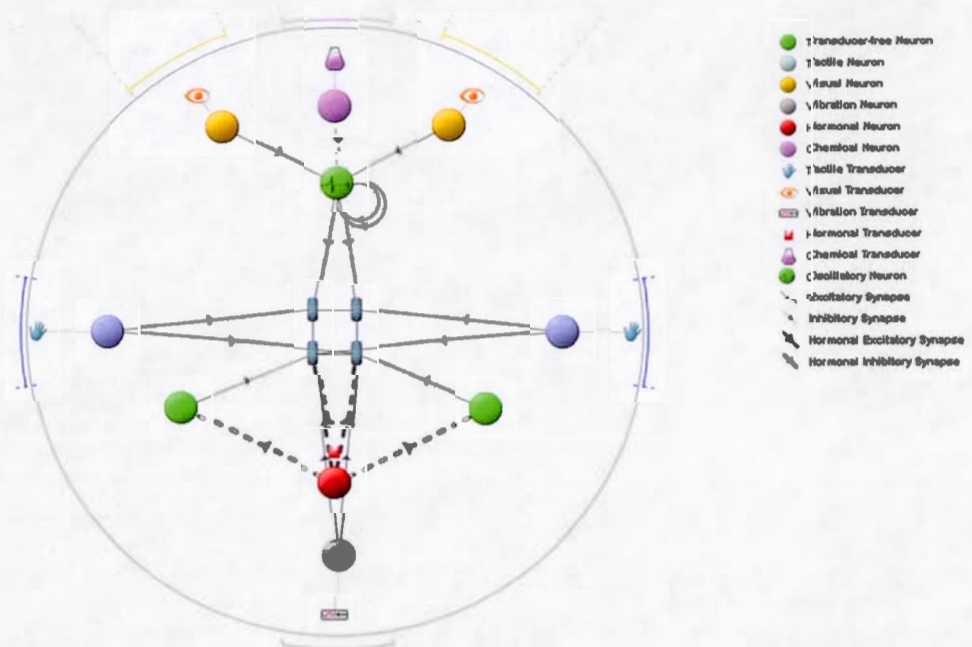
attachments (wheel or turret, muscle remains to be implemented), the spatial configuration implementation on the entity and also on the platform it is associated with, real or virtual. In the virtual platform, movement generation is straightforward and is not concerned with real world factors such as friction. Movement speed is freely set by the user, between 1 and 100% of the built-in initial parameters. It can be linked or not with an energy sensor (battery), giving more realistic robot conditions. With the physical platform, motor adjustments may be necessary depending on the particular setup in use.

Transducers are an unusual component in spiking neuron simulators, which usually accept a file input, often ignoring the holistic view of the artificial life form. In AI-SIMCOG, a transducer is a particular case of a synapse that reflects both the outside view and the inside effect on neurons. The transducer represents the sensory part element attached to a neuron that gives the sensory-neuron receptor appellation. The transducer's main function is to change one physical stimulus into another modality, which is usually an electrical signal subject to modulation in various ways, as it is found in some learning mechanisms. They represent the first contact with the outside stimuli [38]. We think that modelling transducers as integral components of a neural architecture is a more comprehensive way to design complete animats and robots with a biological inspiration. Various types of transducers are offered in the software. They are mostly paired with robotic sensor components for direct code transfer in the physical device. While, sight, touch, vibration and substance detection (for experiment dealing with sort of food

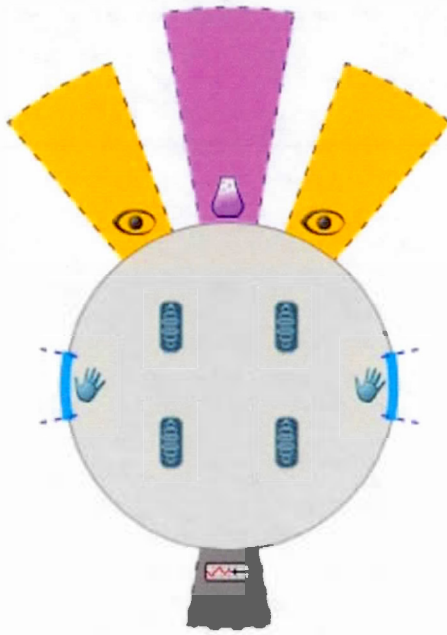
or seeking a chemical signature) are the physical sensory modalities implemented in external transducers, hormones and energy sensors are the internal transducers. The reading of these stimuli receptor devices is made as percentages of full scale. In the case of hormones, a neuron that handles this special type of transducer may target any other neurons with a standard synapse, without any spatial aspect concern. Then, the neuron will pulse and emit signals depending on internal states. The current implementation looks for a kind of stimulation function, similar to the cortisol hormone in natural organisms.

Infrared, ultra-sound, sound, magnetic, thermal and gyroscope sensors are planned to be implemented soon. Transducer sensitivity, field range (width and depth) and type (excitation or inhibition) are editable parameters and the user may connect several transducers to a single neuron, if biologically plausible, with the aim of enhancing topological complexity (e.g. Mexican-hat-type connectivity within a cell). Infinite architectures are possible depending on user needs (Fig. 7). Also, the transducer position on the structure matters, because it reflects the exact position and angle of view in the virtual or physical world. The designer can observe the modification at the animat's outside structure at any time with a special viewer option or animat rendering (Fig. 8). There is also a grid and compass option that helps the visual design of the neural architecture. Print screens in JPEG format of the neural architecture, rendering of the external structure of the animat, weight synaptic table, lists of neurons, transducers and motors with their initial parameters can be made at any time during the edition with a button in the menu.

**Fig. 7** Example of a drag-and-drop, user defined SNN

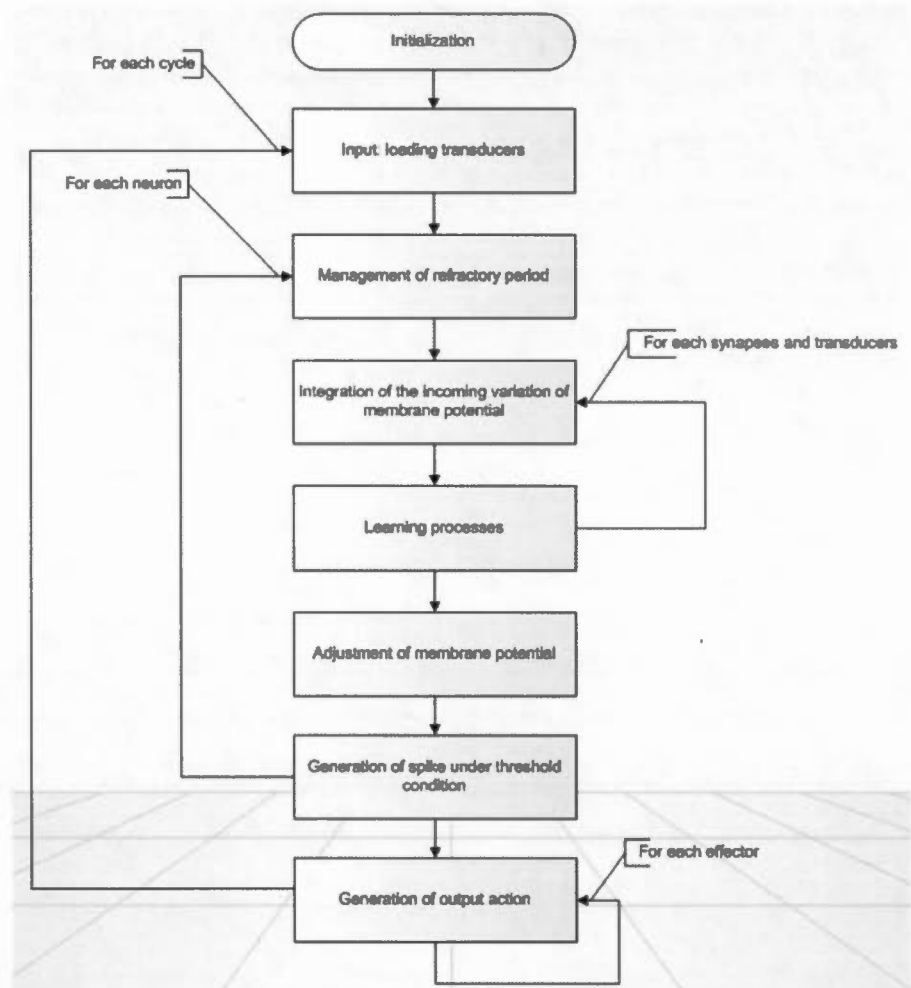






**Fig. 8** Example of a drag-and-drop, user defined animat

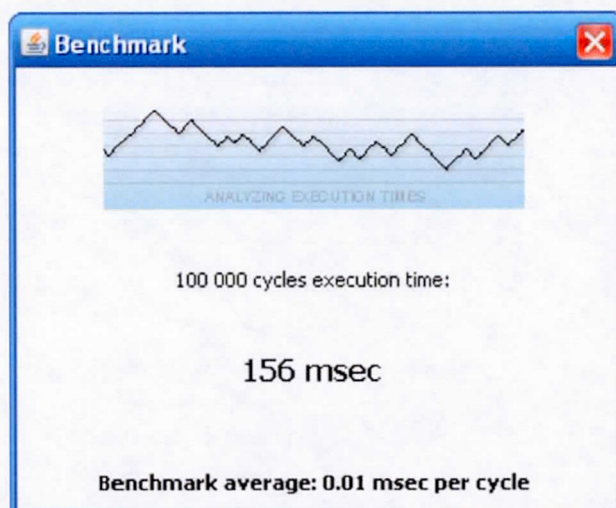
**Fig. 9** Spiking neuron algorithm



## 2.2 Spiking neural network: algorithm

The algorithm behind the SNN model is illustrated in Fig. 9. The corresponding code is generated from templates with all functions and parameters that are editable by the user. Each cycle is periodic and repeats according to a real-time clock, which, in our case, is below 1 ms for moderately complex neural architectures and can be handled by a standard Pentium grade microprocessor. The input–integrate–output processing steps of natural neurons are respected and batch processing is used to generate the SNN outputs. The software provides a benchmark for speed measurement and is illustrated in Fig. 10.

The input function reads and accesses the outside world from the relative positions of the transducers–sensors of all animats. Here also, a transformation matrix is used to convert artificial physical modalities into a variation of membrane potential. Then, the integration of all incoming stimuli (from transducers and other neurons) is done with a subsequent transformation of the plastic synapses, depending on the applied set of learning rules. The output



**Fig. 10** Real time speed for the execution of a simple neural architecture

functions generate spikes, followed by a variation of current in the post-synaptic element. If this element is an effector, an action will follow and be performed virtually or physically if called for by the frequency, intensity and time latency of the input excitation.

### 2.3 Virtual World editor/simulator

AI-SIMCOG also implements a minimalist 2D artificial world, which is sufficient to embed a SNN model in a complete virtual animat's environment. Before starting a new simulation, an artificial world project must be created (Fig. 11). Initially, a flat closed world of programmable size is presented with no objects other than external walls<sup>3</sup>. Again, an icon-based menu allows the selection of items to be added. In the tactile mode, the designer can introduce many different small solid blocks of round and square shapes and can create custom landscapes and labyrinths with them. To speed up application development and to standardise experiments in the learning and memory domains, spatial built-in setups like T-maze and Radial-maze are already available (Fig. 12). All these tactile objects can be edited with a name, size, weight and coordinates. The weight feature finds a correspondence in the transformation matrix that is linked to the animat's tactile transducers—sensors. A vibration feature can also be linked as a special property of some elements such as the tactile one. It uses an array of multiple parameters (number of blocks, time delay between blocks, number of wave emissions within a block, intensity, decay functions and radius) leading to a wide range of output patterns production. Furthermore, the user can create a useful object called

the “Substance” from the left panel. This element serves to simulate a patch of arbitrary substance (e.g. food, chemical), with fully editable parameters such as radius, density, multiple temporal delays and also a special variable concerning the rate depletion of the substance with or without an animat interact with it or not. This item allows the user to build interesting cognitive scenario of navigation, cooperation and competition.

In the light mode, it is possible to choose between two light sources: a spotlight with possible distance and intensity decaying values or a diffused full light such as the sun or a neon light. The same editable features, including dynamical paths as previously described, are available except for the irrelevant weight feature, which is replaced by light intensity.

For each element (tactile, light and substance), the user is also able to generate movement by creating complex paths. This is achieved by clicking on the appropriate path anchors. The dynamics of the moving elements can then be observed during the simulation run.

All the properties mentioned above lead to more realistic simulations where the dynamic interaction with the objects is free to explore and where behaviours beyond standard obstacle avoidance or wall-following tasks may be investigated. It is also possible to explore or capture the relation between relative object speed and time.

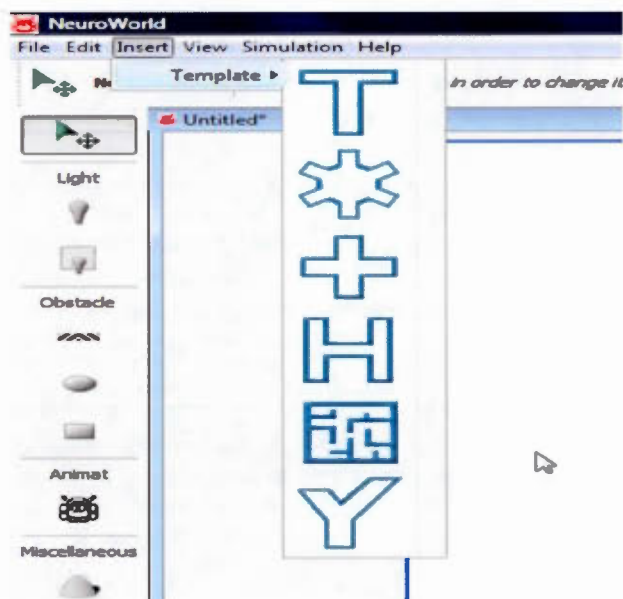
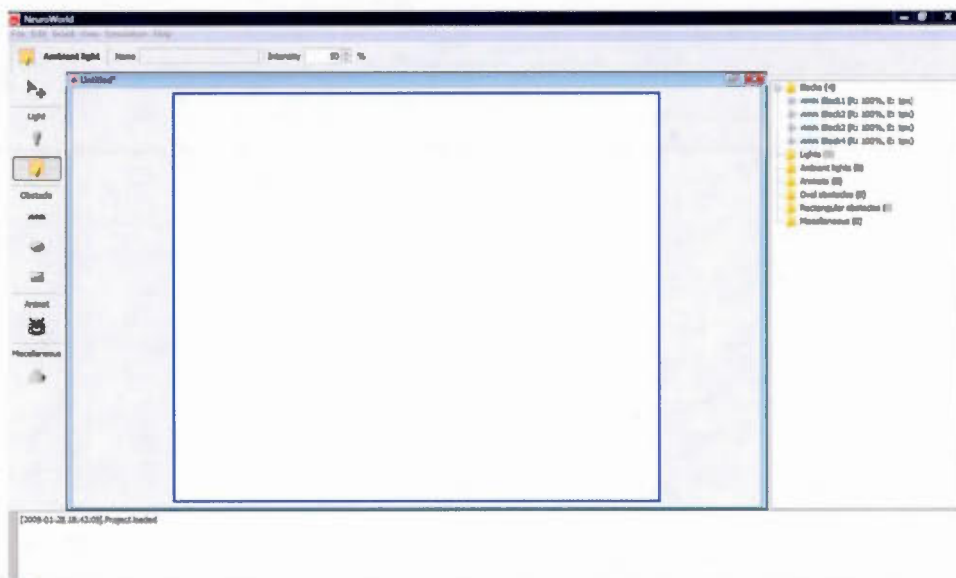
The animat is at the heart of AI-SIMCOG and the user may pick one or more with their own neural architectures. Collective intelligence and emergent social behaviours are common reasons for introducing multiple animats. Size, orientation and initial coordinates in the virtual world are free to experiment as well. At the top menu, the software provides choices to show or hide a compass, a helping grid, object names and a window to remember the outside or internal neural structure of a particular animat. A toolbar is also available to open a stored time-events window at the bottom of the screen.

Once the simulation button is hit, the animats start moving according to their sensor inputs and defector outputs, under SNN control (Fig. 13 provides an example screenshot). Currently, up to ten relatively complex animats can run without affecting the real-time response of the simulator, which smoothly displays the displacement and action dynamics in the virtual world. A dedicated button is available to decrease or increase the speed of the simulation for selected parts. Another one serves for the production of animat visual path trace. Again, screenshots can be accessible at any time. Save, replay scenarios and offline display running are also possible for all configurations.

Data production quickly becomes huge with longer time simulations, complex neural architecture and number of animats. To minimise the memory loading, an option is

<sup>3</sup> An open world version of the software is in progress.



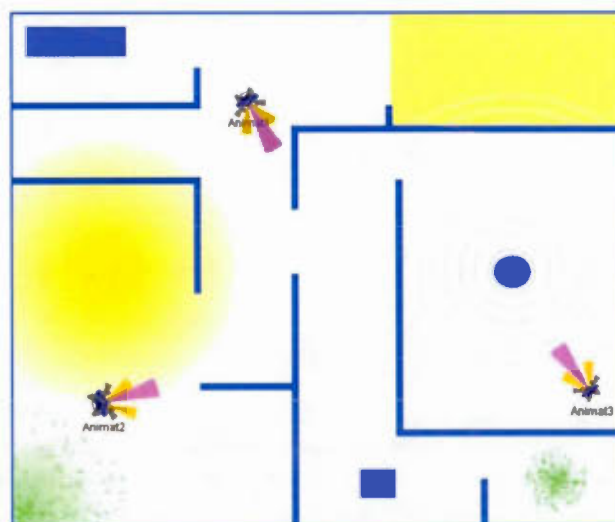
**Fig. 11** Main window of the virtual world editor**Fig. 12** Available built-in mazes

available to save the desired selective variables in the log file.

## 2.4 Software design

This section presents the internal aspects of AI-SIMCOG from the software developer's perspective, i.e. from a design-oriented standpoint, using the unified modelling language (UML) formalism and including a use case and two class diagrams.

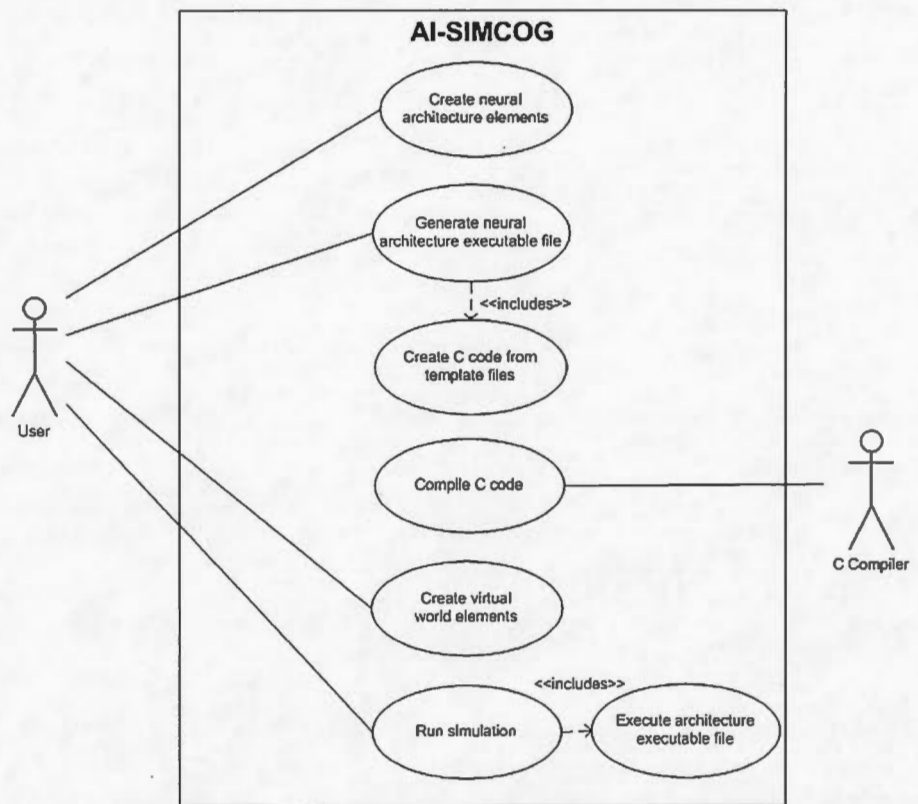
The creation and constant availability of a low-level C language executable was the primary concern of AI-SIMCOG's design, because of the required capability for

**Fig. 13** Example of a custom virtual world designed with animats and various obstacles

physical implementation in robots with small computational capacities and affordable components. Moreover, the main core algorithm for the spiking neural controller was developed to satisfy real-time concerns, namely the millisecond cycle limit that was fixed for biological realism in the temporal and cognitive domains. This led to implement the GUI editor for neural architectural design and the conversion of the resulting project to an efficient executable, readable in both virtual and physical platform modes, as nearly distinct components of the software.

Because the GUIs and graphic components are more easily developed with an object-oriented language and only the executable code needs to be written in C, the software was written in Java with the C code generated via Java

Fig. 14 Use case diagram



scripts or templates with embedded C code sections. Java was chosen to code the GUIs of NeuroCode and NeuroWorld, thanks to its graphical capabilities using Swing, and its portability across platforms. With this approach, it was possible to develop the software both at the level of the inner “science” algorithm and at the level of GUI for project creation and needed graphical facilities. This allowed the software to evolve in both aspects with no interference and to transparently adapt to new C code generation in case of implementing a neural architecture in a new physical platform; after developing and embedding the newer elements’ logic within the appropriate java scripts, the stored executable C code is easily recalled by the main program. In practice, a user-defined template containing both Java and C code produces the final executable file.

Since the logic of the components used by NeuroCode often requires tuning during the software development cycle, the use of external template files or scripts prevents the need to recompile the Java code after every modification made to the components. In addition, if the C language code is migrated to another language, the only files that need to be changed are the template files. Conversely, if the main Java code is switched to another object-oriented language such as C++ or C#, the C code inside the template files does not need to be modified.

As for NeuroWorld applications, the virtual world runs under a Java environment, but using the produced C executable file of the neural architecture. In case of a simulation that features one or more animats, the Java code that is produced after project creation reads from and closes the specified C executable code link to the neural architecture for each animat during the simulation run.

Figures 14, 15 and 16 provide an overview of AI-SIMCOG’s architecture. Figure 14 shows a typical scenario while Fig 15 and 16 show the class diagrams of NeuroCode and NeuroWorld, respectively.

A typical scenario involving AI-SIMCOG begins when a user opens the NeuroCode package with the goal to create a spiking neural architecture. The user chooses items (neuron, synapse, transducer, etc.) from the left menu to achieve his project. After compilation of the code and a quick save of the final architecture, the user may then choose to export his C code to a physical robot or open the NeuroWorld package to design a virtual world. In the latter case, the user chooses again various objects to produce a desire world, including one or more animats that are linked to the previously built neural architecture. The next step for the user is to run the simulation and to observe the evolution of the project and examine all the data and graphics generated by the simulation. The user is then free to vary parameters to support his hypothesis.



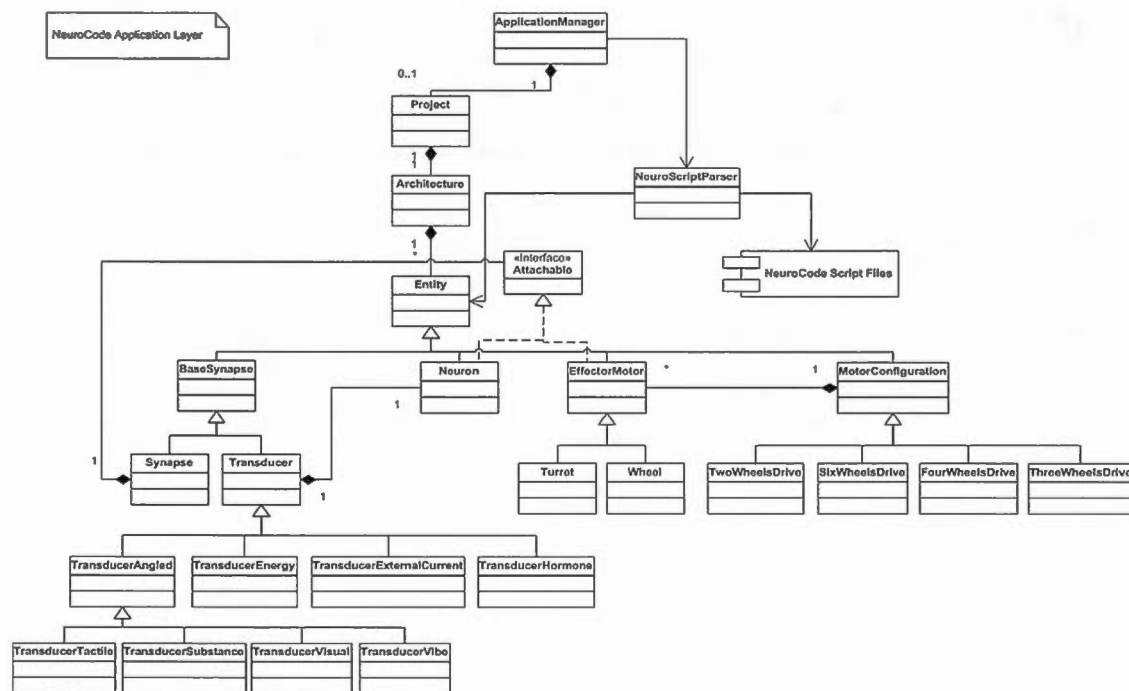


Fig. 15 Class diagram of NeuroCode

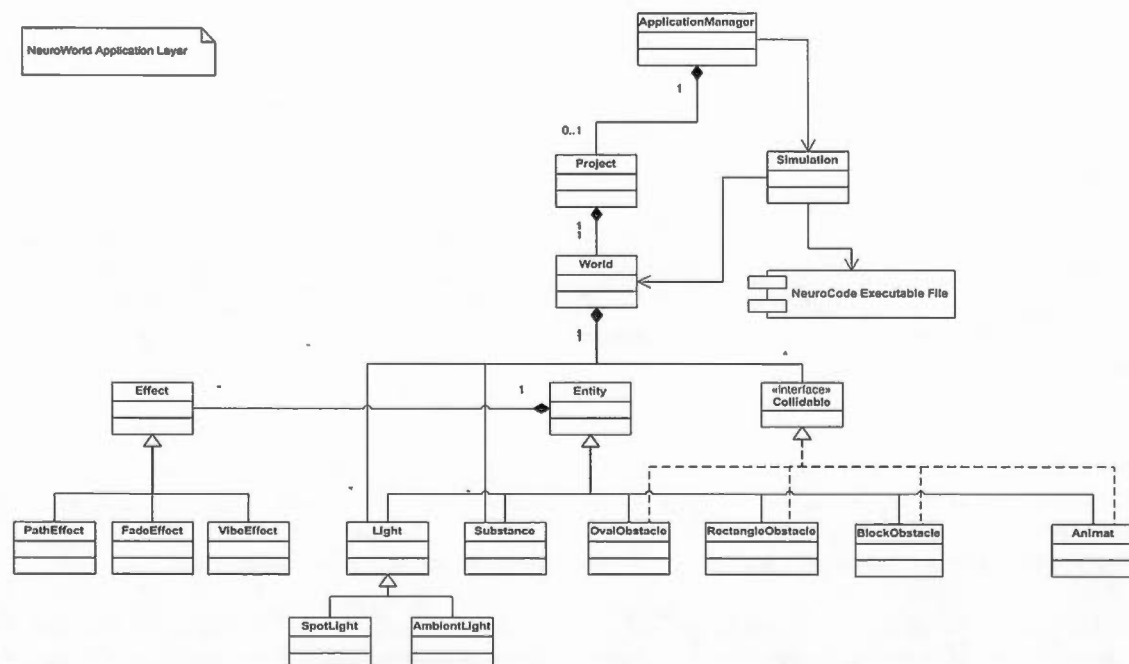


Fig. 16 Class diagram of NeuroWorld

### 3 Application

In order to demonstrate some of the possibilities offered by the software, we created the sample neural architecture shown in Fig. 17. The animat simply behaves in an empty bounded artificial world. What we are looking for here is to

expose graphically the neural fire pattern time sequence in correlation with effector discharges and stimuli events. The story will eventually tell us that the animat adapts in a certain way to its environment. This animat example has features and is hand wired to be just a little more intelligent than a reactive animat. Also, only the habituation learning





Name	Strength	Type
ON/M-RF	90	Excitatory
ON/M-LF	90	Excitatory
TN/M-RR	40	Excitatory
TN/M-LR	20	Excitatory
TS/M-RF	40	Excitatory
TS/M-LF	20	Excitatory
TW/M-LF	10	Excitatory
TW/M-RR	10	Excitatory
TE/M-RF	10	Excitatory
TE/M-LR	10	Excitatory

Fig. 18 Synaptic weight connections table

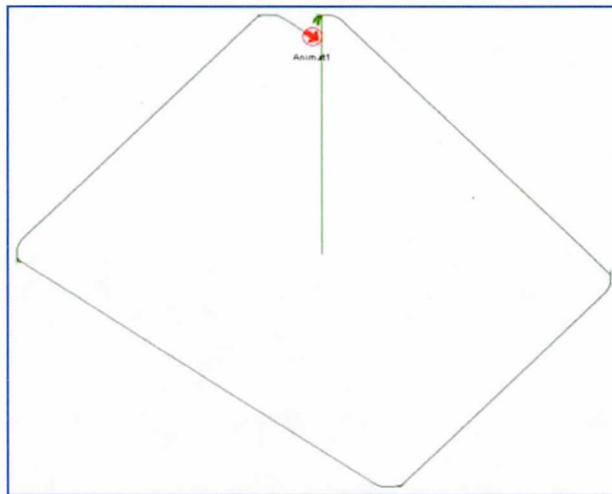


Fig. 19 Path trace of the animat

obstacle in front of it, with a slightly fixed rotation. At this stage, with a very simple neural architecture, the reactive animat will move indefinitely in the empty bounded world with a predictable pattern, reflected in Fig. 19.

These behaviours discussed previously are correlated in the different temporal neural patterns shown in Fig. 20, where ON exhibit a regular spiking pattern every 10 cycles of the algorithm, M-LF and M-RF respond at the same pace with the reception of small post-synaptic potentials. The first spike of TN indicates a collision with the wall ahead of the animat exactly at cycle 130. Soon after, backward movement is provoked by the TN excitatory post-synaptic potential on M-LR and M-RR. These are muted latter for a couple of cycles, while M-LF and M-RF are continuously stimulated by ON, giving another bump on the wall. This situation will only end with several mild turns of the animat, because of the asymmetrical increase strength connection on one side over the other. That is why, NW starts to spike near cycle 200. One can observe in the graphic the temporal and non-linear summation of the EPSE on M-LF and M-RR coming from neurons TN and

TW. With this brief increase of positive current on these effectors, the animat is able to move away from the wall and take another direction with a certain angle until it hits the wall again. Because of this, spikes are evoked in TN and TW at almost the same time for the next bumps. Another evident observation in the graphic regarding the neural architecture example is that there is no spike at all in TE and TS, therefore, no need for them in the simulation.

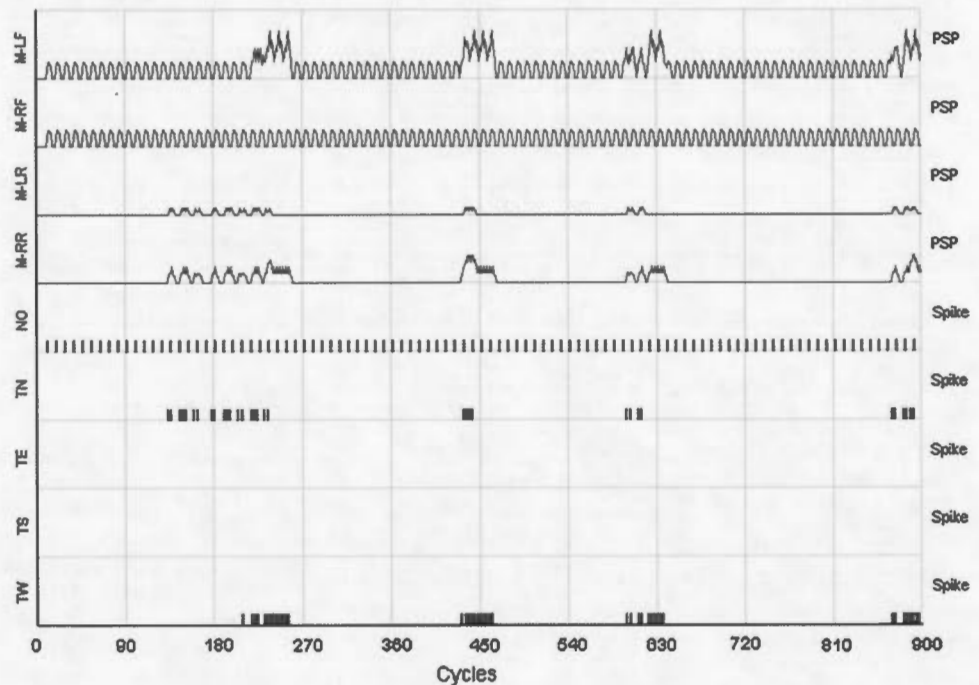
There is also another way to increase the adaptation of the animat without adding other structural and permanent modifications to the architecture. It relies on the well-known habituation neural mechanism. What happens is that with repeated stimuli, the neural response should decrease at the sensory-receptors or between neurons at the synapses. This response may be quite complex. If we examine the situation in our previous simulation, at the first front bump, the reflex to move backwards will not be as instantaneous. Contact with the wall lasts 41 cycles during which the animat operates a turn and then moves away from the obstacle. It happens this way every time the animat encounters an obstacle because of all the parameter settings and kinetic neural curves. With the habituation function (simulation not shown) and because of the specific set-up, the wall contact duration will be slightly shorter with 39 cycles. Also, the input stimuli amplitude overall is decreased in the simulation. Thus, the sensory-receptor model as a synapse, expresses habituation. Moreover, the same run including five sequential bumps against walls is shorter in length of cycles and showed differences in path trace. The interpretation is that the constant ON drive forward behaviour is gaining over TN and the following response of backward movement which decreases with the habituation function.

#### 4 Discussion

Exploring cognitive temporal problems in artificial intelligence was the primary reason for developing AI-SIMCOG. This software offers a multilevel approach, with an emphasis on investigating learning and memory functions at the dynamic and plastic synaptic level and, at the same time, the observations of emergent behaviours of one or more animats in cognitive setup experiments. Validating biological data or developing theories about natural intelligence and brain functions are obvious applications of AI-SIMCOG. Testing hypotheses about the creation of intelligence artefacts in a virtual world or with real robots represents the main avenue, which, with this software, can be of neurophysiologic inspiration.

In our conception of a global intelligence phenomenon, we believe that there is a possible link between natural and artificial intelligence. This bridge may come through a

**Fig. 20** Temporal EPSP and spikes



dynamical integration of real world stimuli in a physical entity. In a biological organism, this process needs a fundamental unit called transducer, which transforms the physical stimuli from the outside into another modality. In most cases, these changes pass through electrical signals. In neural organisms, the stimuli converge at different time scales in a complex neural network to procure a rough sensation of the moment. This sensation represents a partial comprehension of the real world and seems to be enough to act upon it. Wired by evolution, genetics and learning functions, the organisms adapt to their environment by active loops of perception and action. This phenomenon, seen by an external observer is what many researchers call intelligence.

Capturing the essence of these key factors in the intelligence phenomenon is one of the goals of the studies in the field of artificial intelligence. We think that type and diversity, intensity and frequency at various temporal scale levels of input stimuli, integrated in an embodied artefact that is developed from a coherent model of brain-body-environment are some of these main key factors to produce artificial intelligence.

In our personal opinion, we believe in a biologically inspired road with a SNN controller for virtual or physical artefacts. Natural neurons are top integrators and mimicking their functions is a path worth following. With SNN, the temporal domain is accounted for and with added features like dynamical synapses management and basic learning functions, we step ahead in the direction of creating an intelligent artefact in a pure sense. When equipped

with these internal features, a sensitive and reactive physical device may generate active stimuli. Thus, artificial intelligence can be created by transferring natural functionalities to an artificial creature.

In this view of AI, AI-SIMCOG represents a very useful development tool. It is a software that can test and handle many virtual and physical scenarios using SNN computational power for all cognitive problems related to the exact timing of animat behaviours, neurons and synaptic adaptations. That is partially due to the non-associative habituation and sensitisation, and associative short/long-term learning memory mechanisms that are implemented. At present, these functions are at the alpha stage and need further testing and adjustments to be fully validated and verified. The ultimate goal is to demonstrate a strong relation between the extraction and retention of non-linear informational events in different time scales and the corresponding animat or physical robot behavioural adaptations. One can imagine a situation where the two co-evolve in real-time, online and in a life-long situation to procure cues about the nature of intelligence, in particular, the link between all aspects of temporality and the global intelligence phenomenon. For example, imagine a T-shaped maze environment where a light cue, located horizontally near the junction of the T, lights up when an animat that depends on light for its energy balance is close by. The animat could use an anticipative neural associative mechanism to progressively adapt its locomotion guided by the light source. This eventually will lead to locomotion behaviour that is relative to an incoming light energy instead of simply choosing a random pattern to



explore. Of course, a change in the light source parameters and/or location should challenge the learning process of the animat's SNN synapses, which should learn the new time events' association once again. Moreover, a dynamical light cue that is timed or intermittent according to some temporal pattern should also challenge the temporal aspect of the synapse's plasticity, which is one of the main goal applications of the software.

Another original application of this software is in robotics (forthcoming article). In this field, real-time and real-world represent strong constraints where AI must evolve. One of the problems that students, researchers and developers in the field face today is the small number of SNN simulators available that can be linked to artificial life forms at the behavioural level. Some simulators are too general or too specific, others are difficult to understand or use, with poorly designed GUIs, while others simply lack the parameters or special functions needed. We think that our software provides a good starting framework and benchmark tool to fill this gap in AI domain.

AI-SIMCOG also has limitations. Its small network size emulation reflects computational limits that may only be addressed with Moore law or the arrival of massive parallel computers. Also, personal design of a neural architecture may be a very difficult task. Hand-crafted connections between neurons, intuitive synaptic weights and coherent adjustments of huge amounts of parameters are certainly laborious in a relatively big neural architecture. Following repercussions of all temporal variables on the animat's behaviour, may not be easy (or desired) for all users. That is why it is possible to include in the future, random choice functions, genetic and evolutionary algorithms to more "naturally" find the optimal set-up for specific behaviours in targets. On the other hand, we stand to lose the low-level explanations of the behaviours and replace it with a set of other variables and criteria. Another limitation common to all virtual simulators is that intuitive physical transfer functions of virtual worlds may not be enough to get the real insight. Simulated worlds are possibly a dead-end, with actual physical robots being the only valuable avenue to create credible artificial life-forms.

## 5 Conclusion

AI-SIMCOG is a new software tool for the AI community. It aims to help explore the link between temporality and intelligence. Specifically, the software offers a convivial GUI to design neural architectures, and an easy link between a spiking neuron algorithm and a virtual world platform. The biological plausibility and computational

efficiency of the former offers also a direct path for final robotic implementations. Unique features included in the spiking neuron algorithm, such as transducers, hormone components and extended learning rules may confer a clear value to our software for the study of cognitive temporal problems at various levels of analysis.

**Acknowledgment** This software was developed with the help of AIFUTURE in collaboration with Objectif8, two Quebec-based software companies.

## References

- Pfeifer R, Scheier C (1999) *Understanding intelligence*. MIT Press, Cambridge
- Langton CG (1989) *Artificial life*. In: Langton CG et al (eds) *Artificial life*. Addison Wesley, CA, pp 1–48
- Mataric M (2007) *The robotics primer*. MIT Press, Cambridge
- Chiel HJ, Beer RD (1997) The brain has a body: adaptive behavior emerges from interactions of nervous system, body and environment. *Trends Neurosci* 20:553–557. doi:10.1016/S0166-2236(97)01149-1
- Christley R (2003) *Embodied artificial intelligence*. *Artif Intell* 149:131–150. doi:10.1016/S0004-3702(03)00055-9
- Pfeifer R (2007) *How the body shapes the way we think: a new view of intelligence*. MIT Press, Cambridge
- Ziemke T (2003b) What's that thing called embodiment? In: *Proceedings of the 25th annual meeting of the cognitive science society*. Cognitive Science Society, Boston
- McCulloch WS, Pitts WH (1943) A logical calculus of the ideas immanent in nervous activity. *Bull Math Biophys* 5:115–133. doi:10.1007/BF02478259
- Hodgkin AL, Huxley AF (1952) A quantitative description of ion currents and its applications to conduction Hopfield, J.J. (1982) neural networks and physical systems with emergent collective computational abilities. *Proc Natl Acad Sci USA* 79:2554–2558
- Hebb DO (1949) *The organization of behavior*. Wiley, New York
- Gerstner W, Kistler W (2002) *Spiking neuron models: single neurons, populations, plasticity*. Cambridge University Press, Wulfram and Werner
- Maass W (1997) Networks of spiking neurons: the third generation of neural network models. *Neural Netw* 10(9):1659–1671. doi:10.1016/S0893-6080(97)00011-7
- Maass W, Bishop CM (eds) (1999) *Pulsed neural networks*. MIT Press, Cambridge
- Rieke F, Warland D, van Steveninck R, Bialek W (1997) *Spikes: exploring the neural code*. MIT Press, Cambridge
- Izhikevich EM (2003) Simple model of spiking neurons. *IEEE Trans Neural Netw* 14:1569–1572. doi:10.1109/TNN.2003.820440
- Maass W, Ruf B (1996) The computational power of spiking neurons depends on the shape of the postsynaptic potentials. *NeuroCOLT technical report series*
- Gerstner W, Kempter R, Van Hemmen JL, Wagner H (1996) A neuronal learning rule for sub-millisecond temporal coding. *Nature* 386:76–78. doi:10.1038/383076a0
- Bailey C, Kandel E (1983) *Molecular approaches to the study of short-term and long-term memory*. Functions of the brain. Clarendon Press, Oxford, pp 98–129
- Balaban PM (2002) Cellular mechanisms of behavioral plasticity in terrestrial snail. *Neurosci Biobehav Rev* 26:597–630. doi:10.1016/S0149-7634(02)00022-2

20. Caporale N, Dan Y (2008) Spike timing-dependent plasticity: a hebbian learning rule. *Annu Rev Neurosci* 31:25–46. doi:10.1146/annurev.neuro.31.060407.125639
21. Rudy JW (2008) The neurobiology of learning and memory. Sinauer Associates Inc., Sunderland
22. Steidl S, Rose JK, Rankin CH (2003) Stages of memory in the nematode *Caenorhabditis elegans*. *Behav Cogn Neurosci Rev* 2:3–14. doi:10.1177/1534582303002001001
23. Amygdala Grover M, Koch R. <http://amygdala.sourceforge.net>
24. Catacomb3 Cannon RC. Boston University. <http://www.catacomb.org>
25. Genesis Bower J. California Institute of Technology. <http://genesis-sim.org>
26. Matlab Neural Network Toolbox. The MathWorks Inc. <http://www.mathworks.com>
27. Mcell Bartol T. at. Terry Sejnowski lab, Salt Lake Institute for biological studies and Stiles, J. at Ed Salpeter, Miriam Salpeter lab, Cornell University. <http://www.mcell.cnl.salk.edu>
28. Neuron Hines M, Moore JW, Carnevale NT. Duke University. <http://www.neuron.yale.edu/neuron>
29. Simbrain Yoshimi J. SimBrain development. <http://www.simbrain.net>
30. Snnap Byrne JH. and al., University of Texas Medical School at Houston. <http://snnap.uth.tmc.edu>
31. SNNS and javaNNS, Eberhard K. University of Tübingen. <http://www.ra.cs.uni-tuebingen.de/SNNS>
32. Spikestream Gamez D. <http://spikestream.sourceforge.net>
33. Webots C. <http://cyberbotics.com>
34. Riddle DL, Blumenthal T, Meyer BJ, Priess JR (1997) *C. elegans* II. Cold Spring Harbor Laboratory Press, New York
35. Shepherd GM (2004) The synaptic organization of the brain, 5th edn. Oxford University Press, New York
36. Smythies J (2002) The dynamic neuron: a comprehensive survey of the neurochemical basis of synaptic plasticity. MIT Press, Cambridge
37. Grillner S (1985) Neurobiological bases of rhythmic motor acts in vertebrates. *Science* 228:143–149. doi:10.1126/science.3975635
38. Fain GL (2003) Sensory transduction. Sinauer Associates Inc, Sunderland



## ANNEXE G

### ARTICLE 6 : NEUROSIM

#### CHAPITRE DANS LE LIVRE SUR LES MONDES VIRTUELS



# NeuroSim: a Virtual 3D World to Investigate the Intelligence Phenomenon within the Perspective of Bio-inspired Robotic Agentes

André Cyr, Mounir Boukadoum

Computer Science Department

Université du Québec à Montréal, Montréal, Canada

Frédéric Thériault

CEGEP Vieux-Montréal, Canada

## Abstract

This paper describes NeuroSim, novel and unique simulation software in the artificial intelligence community. Part of the SIMCOG suite, NeuroSim helps to investigate the general intelligence phenomenon using a biologically-inspired approach and available robotics facilities. By importing artificial spiking neural networks acting as “brain” controllers into one or many robots, NeuroSim can track all the details of the produced behaviors in its virtual 3D-world. For this, several views for online observations, and substantial data output, help the users understand their simulations. Beyond the creation of virtual robots and primary forms or maze templates that could be added within the environment, the originality of this software consists above all in the integration of multiple sensory modalities and the application of complex temporal patterns of simulated physical effects for the generation of learning studies.

## 1. Introduction

NeuroSim is a 3D-world simulator and editor component of the recently introduced SIMCOG suite, new software in the artificial intelligence (AI) field (Cyr 2009). Embracing bio-inspired robotic (Floreano 2008) and embodiment (Mataric 1997) paradigms, the aim of SIMCOG is ultimately to create intelligent artifacts controlled by artificial neural “brains”,

exhibiting significant elements of natural cognition, as in conditioned learning (Cyr 2012). These produced AI agents would evolve in virtual worlds within NeuroSim or be embedded in physical robots. The software is designed to test hypotheses on the general intelligence phenomenon (Pfeifer 1999) and explore adaptive behaviors through many levels of comprehension. Since the first presentation of SIMCOG in 2009, major developments have taken place to enhance its functionality, of which the NeuroSim module.

Still in the R&D phase, the current version of NeuroSim is mainly programmed in Java, using external libraries such as JMonkeyEngine<sup>1</sup> for the 3D rendering and ODE4j<sup>2</sup> for the world physics. The goal of NeuroSim is to rapidly develop complex cognitive scenarios and adaptive behaviors with complete AI agents in virtual reality, to implement the controller part after in physical robot platforms.

A specificity of NeuroSim is its link to artificial spiking neural networks (ASNN) (Gerstner1999), a subset of the connexionism paradigm as implemented by neural networks. Claiming to emulate some of the critical features of the natural neural elements, the ASNN goes more deeply into biological realism, computing the smallest quantum of information at the level of a single spike. Since the analogy with the biological brain requires input from sensory-transducers and output to actuators to perceive and act on real the world, ASNN is a promising tool for integration when it is embedded as a “brain” controller for autonomous AI agents (virtual or physical) equipped with sensors and actuators. Merging ASNN and robots is of increasing interest in several areas of AI (Alnajjar 2008, Paolo 2003, Soula 2005, Webb 2003). Bio-inspiration for robotic structures (Hopkins 2009) is also seeing growing interest, but no software appears to currently exist that is specifically dedicated to integrate all these fields.

SIMCOG’s ASNN belong to the leaky integrate-and-fire family model, based on membrane potential fluctuations. It uses processes to allow the nonlinear summation of synaptic delayed current inputs, a parameter to backtrack to the baseline homeostasis neural state, a term to generate spontaneous firing and several synaptic learning rules. A key point of this particular ASNN used in the software is its capability to run efficiently within the NeuroSim virtual 3D-world. As such, the ASNN model relies on descriptive curve lines provided by standard mathematical equations in the

---

<sup>1</sup> JMonkeyEngine: <http://jmonkeyengine.com>

<sup>2</sup> ODE4j: <http://ode4j.sourceforge.net>

domain, but implemented in matrices of scalar numbers instead of solving complex functions online.

Taken separately, other non-exhaustive ASNN editors (i.e. NEURON<sup>3</sup>; Genesis<sup>4</sup>; Neural Network Toolbox<sup>5</sup>) and robotic environments simulators (i.e. Simulink<sup>6</sup>; OpenSim<sup>7</sup>; Microsoft Robotic Studio<sup>8</sup>; Webots<sup>9</sup>) do exist, probably doing best in their respective domain, but pursuing different goals in the perspective of their own field. Merging knowledge requirements of two scientific fields may find advantages in tightly coupled software such as the SIMCOG suite.

Thus, NeuroSim focuses on bio-inspired controllers, insisting on the neural integration of multiple sensory modalities and the observation of concrete adaptive behaviors at the agent level. Emphasis is put on the functional over the mechanical aspects of the AI model. As such, accuracy of simulations regarding the graphical aspects and biological exactitude of the neural processes are not the main targets of NeuroSim. Instead, coherence of the dynamical cognitive system, bridging the controller and the sensors/effectors components, balancing the physical robot's structures and its biological agent's counterpart, and creating intuitive interfaces are the main elements kept in mind during the development of this module.

The first step in using NeuroSim is the creation of an empty 3D-world. Then, the user adds one or more robots and associates them with the appropriate ASNNs and body structures files produced by another module of the SIMCOG suite<sup>10</sup>. For instance, the user can easily explore the effect of heterogeneity of internal and external AI agent's structures on a given environment or task. To achieve this, the user would enrich his/her experimental scenario by including several additional objects available from a toolbox. Ultimately, the goal is to start a simulation, watching robots evolve in the 3D-world, perceiving stimuli and acting in the environment through their sensors and motors. The general functioning of the simulator stands on the repetition of regular discrete steps, consisting in sensory lecture from perceptive animats and generation of actions and stimuli into the environment from the objects, where the animats represent a particular subset of these objects. Temporal and spatial conflicts are managed at best

---

<sup>3</sup> Neuron: <http://www.neuron.yale.edu/neuron>

<sup>4</sup> Genesis: <http://genesis-sim.org>

<sup>5</sup> The MathWorks: <http://www.mathworks.com>

<sup>6</sup> Simulink: <http://www.mathworks.com>

<sup>7</sup> OpenSim: <http://opensimulator.org>

<sup>8</sup> Microsoft: <http://www.microsoft.com/robotics>

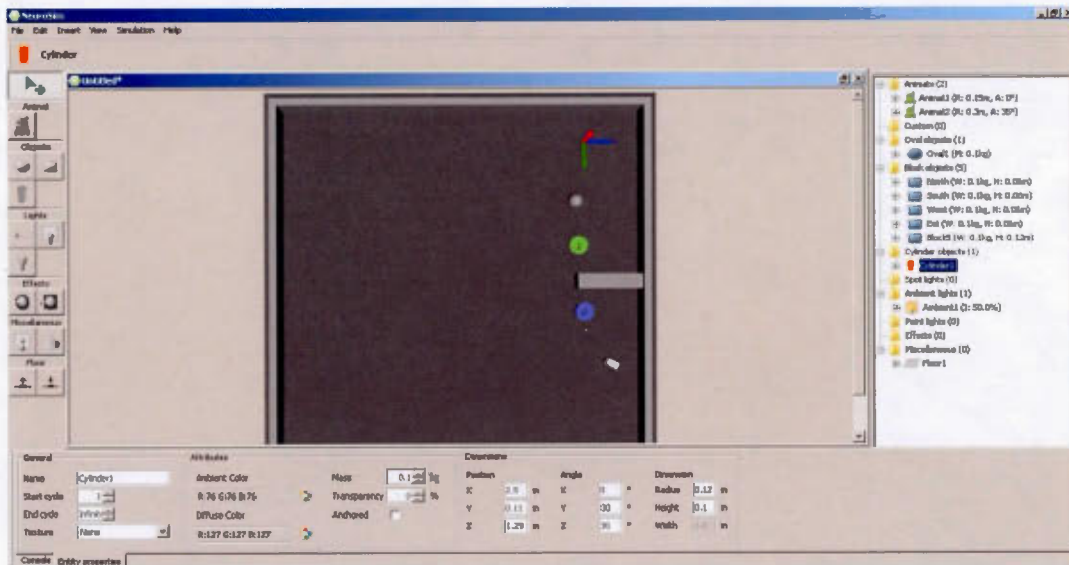
<sup>9</sup> Webots, Cyberbotics: <http://cyberbotics.com>

<sup>10</sup> NeuroCode: <http://www.aifuture.com/neurocode>

to respect the timeline and parallelism of the real-world, composing with the capabilities of the physical engine. This simulator largely puts efforts on the diversity of stimuli produced from the objects into the virtual environment and perceived by robots. Perception is achieved through several sensory modalities and the production of different physical stimuli is done from applying external temporal patterns from the objects or the environment, as it could be retrieved in experimental learning protocols applied to natural agents.

In resume, NeuroSim stands in the middle of neurosciences' current knowledge and robotics research with the aim to develop bio-inspired AI virtual agents exhibiting significant elements of cognition through their adaptive behaviors. This paper describes the graphical-user-interface (GUI) functionalities next, the internal aspects of NeuroSim in section 3 and issues of the NeuroSim module in the Discussion section. We present here the software as a novel contribution in the domain of virtual worlds, when put in conjunction with the bio-inspired robotics field.

## 2. GUI Description



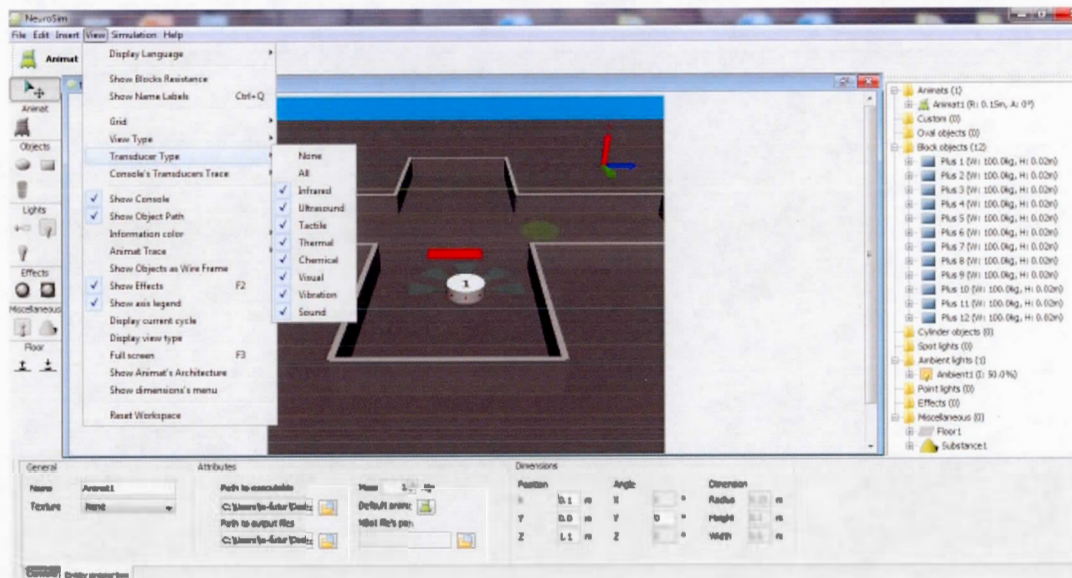
**Figure 1.** The main GUI is composed of a top menu allowing to manage the input/output files and different viewing setting, a left menu used as a toolbox for inserting objects into the virtual 3D-world, a right arborescent menu serving to access all included objects, a central screen for the environment and a bottom pane to unravel either the console menu for actions feedback produced into the simulation or a transient properties menu of a selected entity.



The GUI is simple and very intuitive (Figure 1). Thanks to the presence of many icons and tooltips, the user can build complex environments and simulations with just a few clicks of the mouse, without programming any line of code. The opening window is composed of a main menu at the top, a central screen for viewing the virtual 3D-world, a toolbox menu on the left, an arborescent menu on the right to quickly visualize all items within the simulation and, finally, a console menu at the bottom for control feedback of all actions taken by the software or for pop-up menus showing selected entity properties.

## 2.1. The Main Bar Menu

All essential input-output and internal parameters adjustments are found under the File and Edit buttons. Once a new project is started, several settings concerning the current environment can be adjusted for viewing preference by the user from the View button. Selecting the type of view for observing the simulation (first person, third person, orbital, top or custom), displaying specific kinds of transducer ranges or robot traces are just a few examples of what can be controlled (Figure 2).



**Figure 2.** Snapshot of the main top menu access. As an example with a third person viewing, a user intends to set the visibility for all range transducers from the top menu and the view button; follow by the Transducer Type bar. In this virtual world sample, a typical Plus maze was previously added in one click of the Insert button (also in top menu), maybe for navigation and foraging experiments. A red strip on the floor may also be used as a contextual learning cue to hypothetically tell the robot that there is a substance to catch (painted in green) on the right arm of the Plus maze.

The Simulation section brings the user to a dynamical mode through a new left panel, offering traditional play-stop-pause and single step buttons. Taking pictures or creating basic videos is also an available option in this section. During the simulation, it is possible to pop up small windows to show the current state of a selected robot, visualize the neural architecture or add other viewing windows. Another interesting icon in this panel permits to freeze the virtual world for an online object teleportation, instantly bringing any object (robot included) to new coordinates. Finally, depending on the computer it runs on or the purpose of the experiment, speed control for the cycle time length cycle can be modulated through a sliding bar.

## **2.2. The central screen**

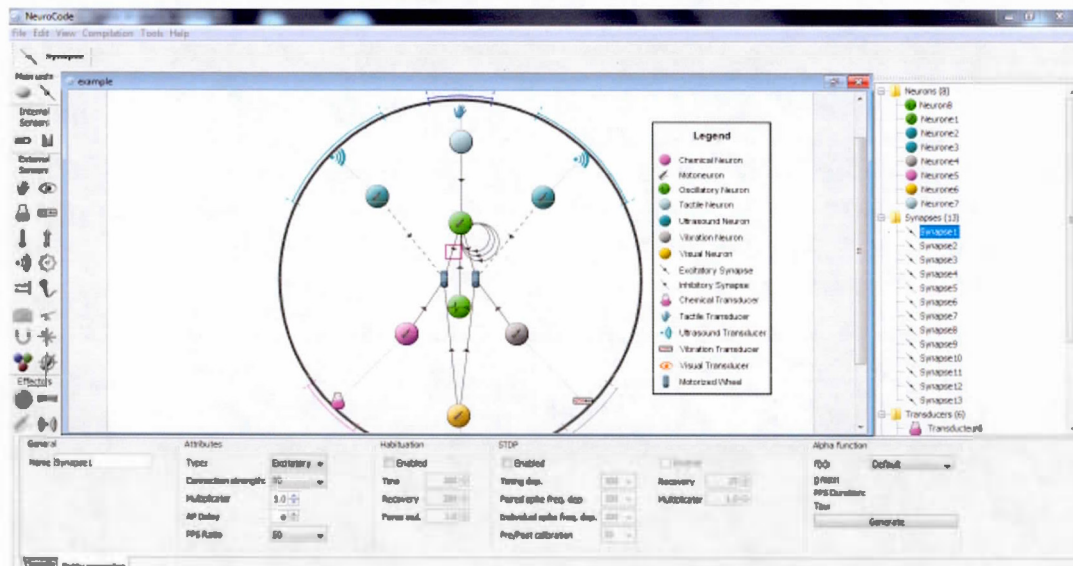
The opening window starts with a closed world of 6 meters x 6 meters that can be set otherwise. In this configuration, four anchored walls delimit the virtual environment, but they can be removed by a simple drag and drop option to create a world with no boundaries. Clicking inside the world allows the user to identify and move any object inside or fill the environment with new objects from the Toolbox left menu. Once an object is selected in the central screen, the attached special properties automatically appear into a popup menu at the bottom of the window. These properties depend of the selected object, but standardization lets the user quickly learn how to adjust them. Non-exhaustively, the characteristics of an object consist in: position using x, y, z coordinates, angles and dimensions (length, height and width). Color, mass, transparency, anchored or not and when it appear or disappear during the simulation are others options ready to set.

## **2.3. The Toolbox Menu**

It is with this menu that a user really creates rich virtual worlds. Choosing to insert a robot is certainly the first step to accomplish tough simulations could be run with without agents. After clicking on the dedicated robot icon, the user can point anywhere in the central screen, and, with a left mouse button click, the robot is deposited on the floor. Then, left clicking on the robot again allows temporarily leaving the toolbox menu and making the transient menu appear at the bottom, where the robot properties are available. As a special object, a robot (agent/animat) needs to be associated with a valid input/output controller to evolve in the virtual 3D-world: an ASNN file consisting of a neural architecture controller incoming from NeuroCode (Figure 3) and an output file path name to write all produced

data. Finally, several different robots of different shapes and neural architectures could simultaneously evolve in the same simulation.

Below the animat button in the toolbox menu (figure 2), basic forms are available for selection, enriching the virtual environment. Blocks of any size are useful to create obstacles or object to be pushed or pulled. Beyond blocks, ovals and cylinders are the other primary objects that could be inserted into the world. Aggregating several of these simple forms of possibly different dimensions, colors, textures and several other properties could eventually lead to building complex structures. Also, three types of light are available in the toolbox menu. Point lights are directed beams, ambient lights are used for general illumination and spot light represent a kind of directed theater lights. For this particular sensorial modality, the color, intensity adjustments and fading or constant light dispersion are special properties that are settable.

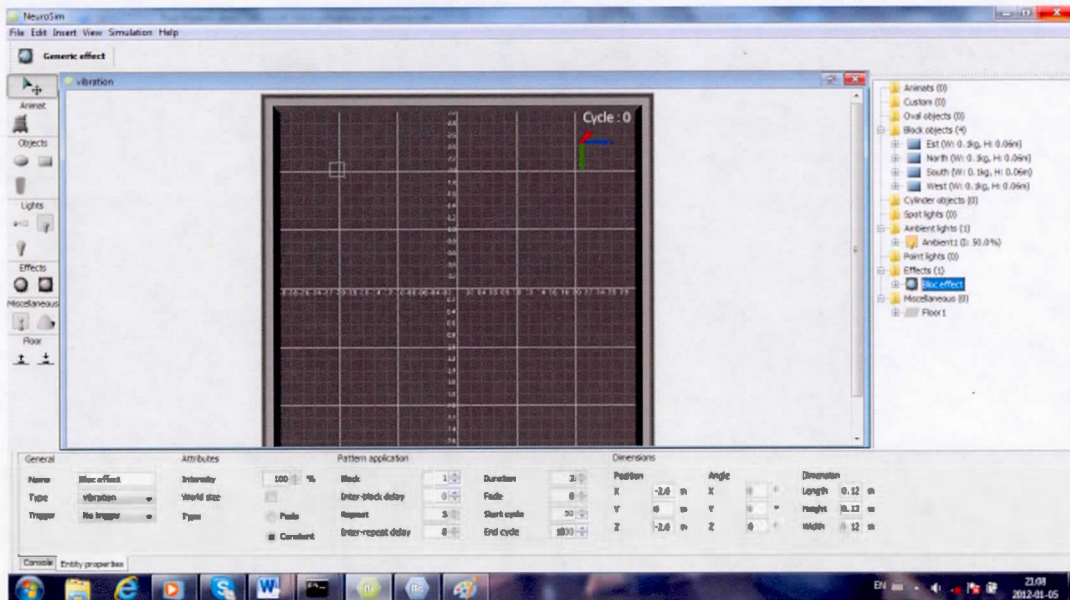


**Figure 3.** Snapshot of the NeuroCode GUI, the ASNN editor of SIMCOG. Users build brain controllers for AI agents, whether running in the NeuroSim virtual environment or physical platforms. In this example, several types of neurons were created with various sensory modalities and different synaptic links. The Synapse1 element illustrated in this view demonstrates the setting possibilities for the parameters of this entity at the bottom window.

A hallmark of the NeuroSim module is the possibility to produce virtual effects of different shapes (oval or block) over several modalities (light, thermic, sound, vibration) by selecting the corresponding menu icons and depositing the effect unto the world; clicking again on the chosen item gives instant access to the effect properties. From this feature effect, a user



could create almost any temporal and spatial pattern. As an example, for a vibration stimulus block-effect production (figure 4), a user could apply the following specific pattern: 1 block of 5 repeats with an inter-stimulus interval of 8 cycles, each stimulus lasting for 2 cycles, starting at cycle number fifty and generated the effect at the top left corner of the 3D-world. This vibration wave is set to be static (dynamic path is possible), constant with 100% intensity all over the field effect and the duration of application (fading in time and space with an infinite decrease pattern is possible). The effect is also not triggered by a specific event (i.e. external device, block move in a certain area, switch turn-on from a robot or robot passing into the area). Figure 5 shows the data graph reproducing the stimulus pattern discussed above, when caught by a vibration sensory neuron.



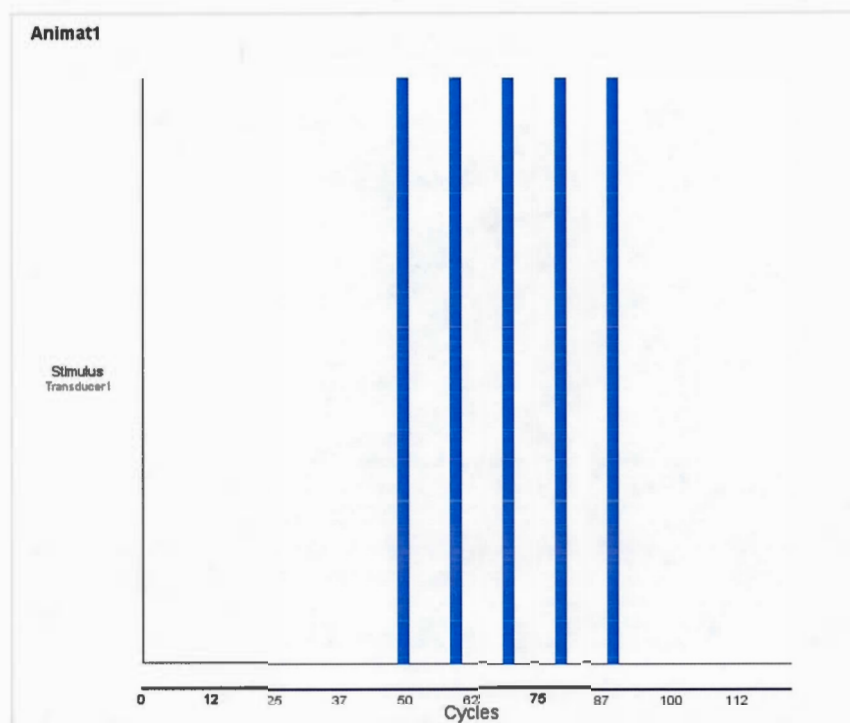
**Figure 4.** Example of physical effect production: The user sets all parameters from the bottom feature effect menu, to simulate a vibration wave effect into the virtual 3D-world, occurring from a precise spatiotemporal pattern (see text) into the experiment protocol. The coordinate's grid tool helps to place the item in the world.

The possibility to combine effects as needed and merging several types of effects with different temporal patterns is convenient to build complex experiments for learning and exploring adaptive behaviors in AI agents. Moreover, the effect pattern generator function is applicable to all objects, of which light objects gives access through the light object properties. Thermic radiation emitting from a heated block source, or emission of a vibration from a cylinder, or sound emission from a robot are just few examples of scenarios that could produce real dynamical interactions



between robots and the virtual environment. At this point of the software development, all modalities are represented as a 3D-map built with discretized percentage numerical scalar values, accessible at every cycle in all spatial points of the world.

Also in the tool menu, under the miscellaneous category, a cube sensor of any size could be added into the world as a special neutral object. Depending on the chosen attached sensory modality, this cube sends a file output of all numerical values it perceives inside, as long as the period is set. This feature is sometimes useful to understand effects in the environment. In the same category, an interesting component is proposed with the Substance. Substances are objects defined by their given names.



**Figure 5.** Bar data graph representing a specific temporal stimulus pattern of vibration waves caught by a vibration sensory neuron, part of a robot's ASSN controller.

Thus, a Substance object could be understood as food, odors, pheromones, etc. Beyond the properties of dimensions, fading or constant dispersion of the Substance, this object also possesses a random dispersion option on different pattern variation. Moreover, substances could be relying on an evaporation mode, washing out the substance among other settable spatial and temporal patterns. Thus, these objects raise the possibility of

creating many different scenarios in learning, foraging and navigation cognitive themes among others.

Finally, the last component of the Toolbox menu is a tile editor and initializer (floor painter and eraser). Coloring some tildes of the floor is standard protocol to observe and generate different robot's behaviors. It is also possible to create mountains or depressions in the floor environment to experiment with gravity, equilibrium, acceleration, etc. This toolbox menu is basic at this stage, but versatile and effective enough to enhance the complexity of the virtual worlds built.

## **2.4. The Arborescence Menu**

On the right side of the GUI, the arborescence menu represents all objects included in the current environment, organised into logic categories. When an object is put into the virtual world, its name is automatically added under the proper category in this menu. Left-clicking on the object in the arborescence menu gives access to its attached properties, with a pop-up menu appearing at the bottom of the GUI. It is helpful to use this viewing mode to reach objects or effects when many of them may possibly overlap in the same area, be hidden by other objects or simply be not visible in the current view.

Right-clicking on the name objects in this menu brings new options to the user. As an example, for a simple block adding, a path and effect buttons are available for this object. Creating a dynamically chosen or random path with different speed for different segments is very simple to add as feature for this new block. Moreover, the dynamical block could at the same time emit sound, vibration or radiate warm from the effect option. Eventual spatial conflicts among objects are resolved with a built-in stop and restart algorithm. Right-clicking on the object in the right menu also brings the possibility to delete it or display the name of the object in the world. Other options are available depending of the selected objects.

## **2.5. The Console**

This menu is situated at the bottom of the GUI and can be closed at any time from the view button at the top main menu. It gives all the sequential information of processes achieved in NeuroSim, such as when objects are created or deleted. It also gives the information about the generation of effects at the precise cycle and the perception of any type and intensity of stimuli from any transducers for all animats. Thus, traces are extensive for almost all variables. Data are output in this window but, it is possible to

filter the amount of information displayed with parameter setting, while saving the complete amount in files for future reading.

## **2.6. NeuroData**

NeuroData is another module of SimCog; it allows creating basic graphs of any variables produced in NeuroCode and NeuroSim. Because there is a constant link between NeuroSim and NeuroData, it is possible to dynamically observe any parameters as graphs during a simulation from an add-on window, for conveniently tracking information (see figure 5).

## **3. Internal Mechanics**

In an effort to be compatible under Mac, Linux and Windows operating system, NeuroSim was developed in Java. Every visual frames and windows are created with Swing, except for the virtual 3D-world.

### **3.1. Creating the Virtual 3D World**

NeuroSim has seen major improvements over the years. After the first few releases of a virtual 2D-world using Java custom drawing, it eventually evolved into a 3D scene. In order to have satisfying performance and all the functionalities required, the game engine JMonkeyEngine (JME) 3.0 was chosen for the 3D rendering. The fact that JME can also be embedded easily in Swing applications is also convenient. One of the continuous challenges with NeuroSim is to stay computationally efficient while expanding the software. Given its real-time and parallelism concerns, as well as the processing of every components of the neural architecture at each cycle, the 3D-world considerably raise the amount data. NeuroSim is not event-driven and for a given simulation every cycle takes approximately the same constant time (except for internal interruptions from the exploitation system in virtual world, but strictly identic in physical platforms). Thus, the more complex the AI agent's controller-brains, the amount of agents and number of objects within the scenario simulation, the longer will be the time period to execute one cycle of algorithm.

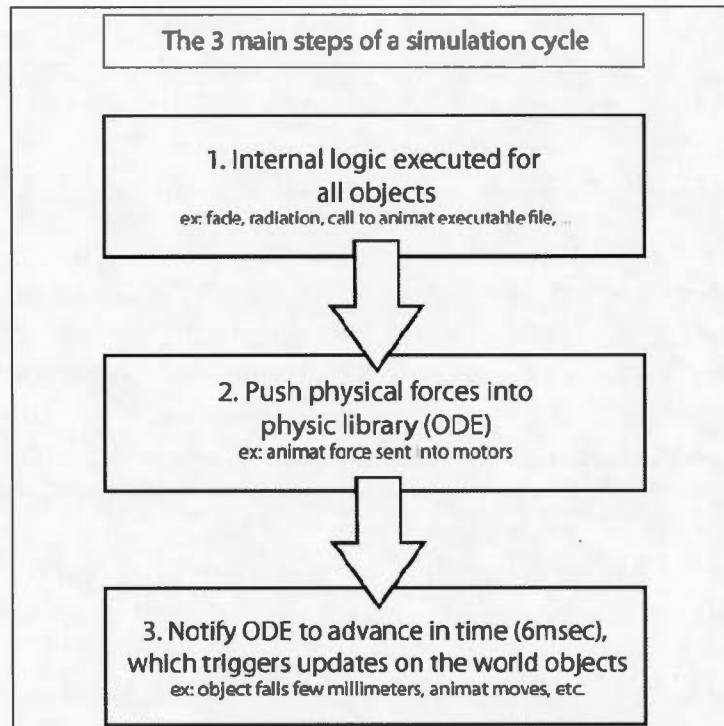
### **3.2. Realistic Simulations**

To be able to create simulations in NeuroSim that mimics reality as much as possible, ODE4j was also embedded. ODE4j is a Java port of the C++ rigid body dynamics library ODE. It provides all physical calculations for objects collisions, gravity, friction, joints and more. This allows NeuroSim to quickly create simulations with moving objects that reacts like real world

objects. For more accuracy, other mathematical tools could be used eventually.

### 3.3. Step by Step Simulations

When pressing the start/play button in the simulation panel, cycles are completed until the simulation ends or is manually stopped or paused. A cycle consists of the following steps.



**Figure 6.** Logical steps describing a typical cycle of algorithm in NeuroSim.

First, internal logic is executed for all objects. For example, if an ambient light object has to fade in a given time window (i.e. 100 cycles), its light intensity will be reduced by a given function (i.e. linear: 1% at every cycle) under it goes dark. For an agent, the internal logic refers to run the executable code (the brain architecture) produced by NeuroCode. By communicating with the agent executable file, NeuroSim can provide sensor data (vibration intensity, for example) and get to know what motor force to put in each of the agents 3D motors in the virtual world. Once the internal logic of all the objects is completed, the next step is to push the force and physic modifications into ODE4j. Only physical objects are considered in ODE4j. Since a light doesn't really respond to gravity or friction, ODE4j does not know that it exists. Finally, NeuroSim tells ODE4j to advance in



time, which makes the objects in the virtual world move according to their friction, gravity, etc. A visual graph of the main steps of a cycle is shown in figure 6.

### **3.4 Network Server**

Since the behavior of an animat in NeuroSim can sometime be hard to explain due to the complexity of neural architectures, NeuroSim acts as a server and accepts TCP/IP connections from NeuroData. Connecting the two allows the user to see live plot graphs of any software variables during the simulation runs, such as membrane potential of a particular neuron, synaptic variation, number of step forward, etc.

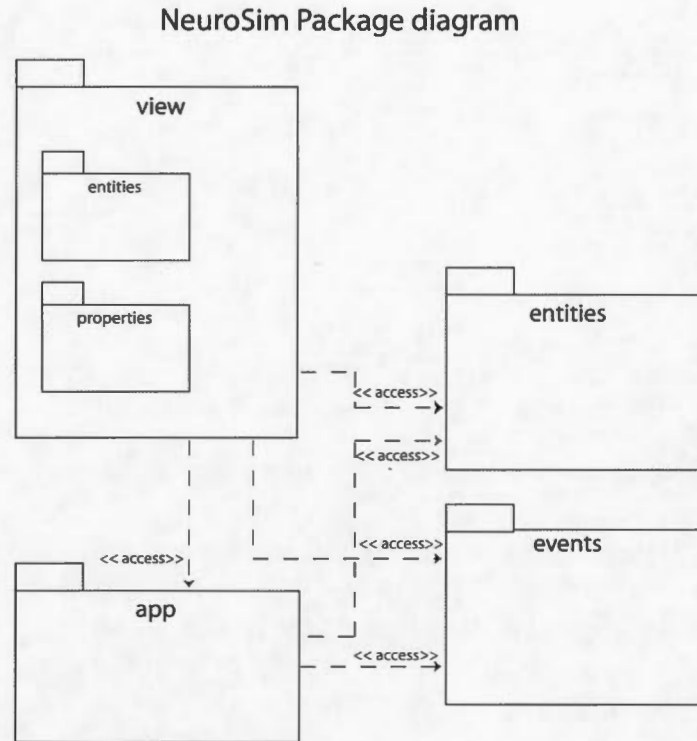
### **3.5. Using Layers to Keep a Clean Development**

Flexibility for further transformation or change in the physical engine is of major consideration in the architectural choice of the development for loose coupling. Therefore, all components are separated from one another, except for the internal Java classes, core of NeuroSim. Reducing coupling between the different layers of NeuroSim allowed, for example, to easily migrating from a virtual 2D-world to a full 3D-world. Some of the major layers are: the physical engine, the 3D rendering logic, the network communication protocol used to communicate with NeuroData, the protocol used in the communication between the NeuroCode executable file and NeuroSim.

NeuroSim contains over 200 classes that are well separated in different packages which allow a logical separation of classes according to their roles. The diagram shown in Figure 7 provides a representation of the main packages, which reflects the internal architecture of the application.

In NeuroSim, all world objects are defined in the entities package. For examples, the Animat class defined in that package will contain all the necessary logic to communicate with its executable file and the Light class object will contain all the logic to be able to do a fade in/out option.

The view package is responsible for the visual aspect of the applications. While it contains other smaller packages (i.e. camera, utilities, prompt, etc.), entities and properties are the main ones. All Swing components are also put under the view package. The property package is used for showing panels, where it is possible to modify the values of each entity's properties. Finally, the entities package under the view package contains classes allowing the visual rendering for each entity into the 3D-world.

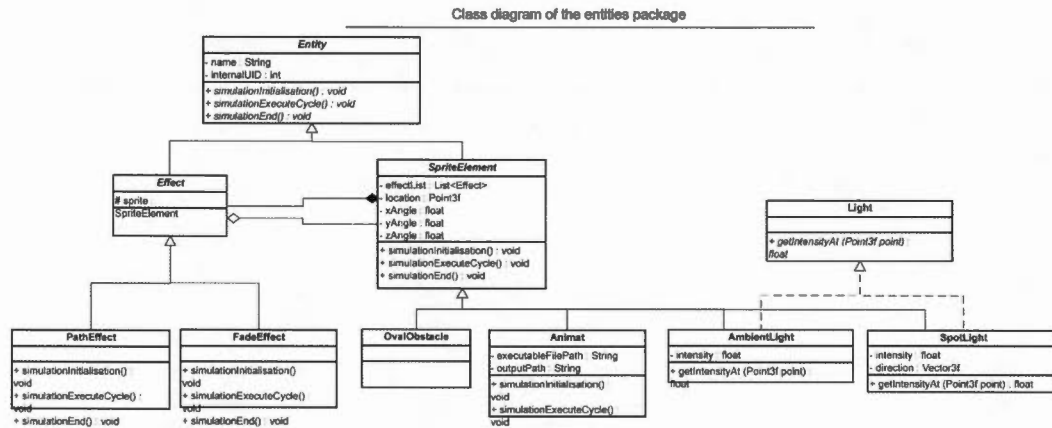


**Figure 7.** Overview of the internal packages organisation of NeuroSim.

The event package contains simple classes that are instantiated by the view package and then sent to the app package. An event is an action that changes a NeuroSim project, triggered by the user. Examples of classes that are defined in that package include `CreateAmbientLightEvent`, `UndoEvent` and `DeleteEntityEvent`.

The app package contains classes that possess the internal state of the applications. The main classes are `World`, `Simulation`, `ApplicationManager` and `Façade`. Using the Façade singleton, all events created by the view are sent to the `ApplicationManager` in order to be processed. Since the internal state of NeuroSim has to update the view, the observer design pattern is used. Therefore, the view package is completely hidden from the app package and it prevents unnecessary coupling. Following is a class diagram of the entities package (figure 8). The diagram does not contain all classes nor do all class attributes and methods, but rather focuses on showing the internal structure of the package.

All NeuroSim objects are derived from the `Entity` class, which obviously contains methods and attributes common to all world objects. `SpriteElement` objects, like blocks or animats, may contain a list of effects that modifies its behavior and appearance.



**Figure 8.** Class diagram representing the main classes found in the entities package.

The coupling between an Effect and its Sprite is quite strong, but it was designed this way for two reasons: an Effect needs to be able to directly modify its Sprite, and the lifecycle of an Effect is decided by the owning Sprite. Related methods are executed during a simulation and allow all entities to update themselves at each cycle.

### 3.6. Communicating with NeuroCode

NeuroSim communicates with the neural architecture using a C executable file instead of other means such as TCP sockets from NeuroCode. The rationale for this use of two programming languages is explained by the ability of the software to produce executable files that can run on multiple physical platforms like the Lego MindStorms NXT (Lego Group) robot or the Khepera III (K-Team Co.). By using C executable files for all robots, it simplifies the complexity of maintenance.

## 4. Discussion and Future Work

NeuroSim is a virtual 3D-world simulator and editor still under development. This paper aimed to describe the functioning of the software in regard of issues in the AI domain. The novelties of introducing this software are founded on merging the dual neuroscience field with ASNN controllers, and the bio-inspired robotics domain with the versatility of structural agent's composition, as well as simulating the high number of perceptive dimensions available into the real world. As such, we unified these paradigms under the NeuroSim module in a virtual 3D-world environment.

The authors' intentions are currently favouring orientation toward academic goals and an open-source product. Since robots as well as

individual sensor and actuator components are increasingly inexpensive and easier to use, the possibility to create agents with “intelligent” behaviors goes now beyond the realm of a small restricted class of scientists. In the past, programing a full logic code for virtual or physical implementation in complete autonomous robots was difficult for most novice users; the user-friendly NeuroSim module now overcomes these traditional difficulties. As an example, without particular knowledge or documentation, a user may in less than five minutes create simple phototropic behavior with few neurons and synapses forming a basic neural architecture created with NeuroCode, and view results through a robot agent in NeuroSim or transfer it to a Lego Mindstorms NXT robot. On another hand, the software could be used by specialists to explore cognition at a high level, merging complex conditioning learning task protocol to low level natural neural synaptic embedded processes as an example scenario. It is a feature of the software to target rapid development of hypotheses on intelligence without programming any code and using instead an intuitive GUI.

We showed in this paper, that NeuroSim is a virtual 3D-world that proposes basic object insertions but high possibilities of interaction from diverse perception modalities into “intelligent” robots. These are the result of ASNN controllers produced from NeuroCode, another module of the SIMCOG suite.

The natural expansions of the NeuroSim module will be to increase the sensory toolbox with additional modalities, be they complex, unusual or non-existing yet in robotics sensors, or natural sensory modalities. For example, merging force sensors of different sensitivities with “pain” detectors and thermal sensors will permit to mimic an “E-skin”, a forthcoming robotic component. Thus, NeuroSim follows current biological researches, waiting for available products in the robotic industry, but already proposing models for virtual autonomous bio-inspired robots.

Our very next move is to develop a CCD-grid camera sensor component (light: extending UV and IR) to enhance the input from the visible spectrum modality. This is currently a challenge given the limited computational resources to create a complex ASNN, especially when it is linked to many sensors and effectors within a rich dynamical environment. Also, since laser beams are commonly used in robotics for environment mapping and navigation, we intend to model soon this sensor (although it has no obvious biological counterpart). The “E-nose” is another complex sensorial modality we plan to implement; we already have the Substance as a regular component of NeuroSim, and volatility and density are just parameters to add. In the end, we have software capable of bidirectional



investigation in parallel: robotics possibilities and currents understanding of biological neural controllers, sensors and effectors.

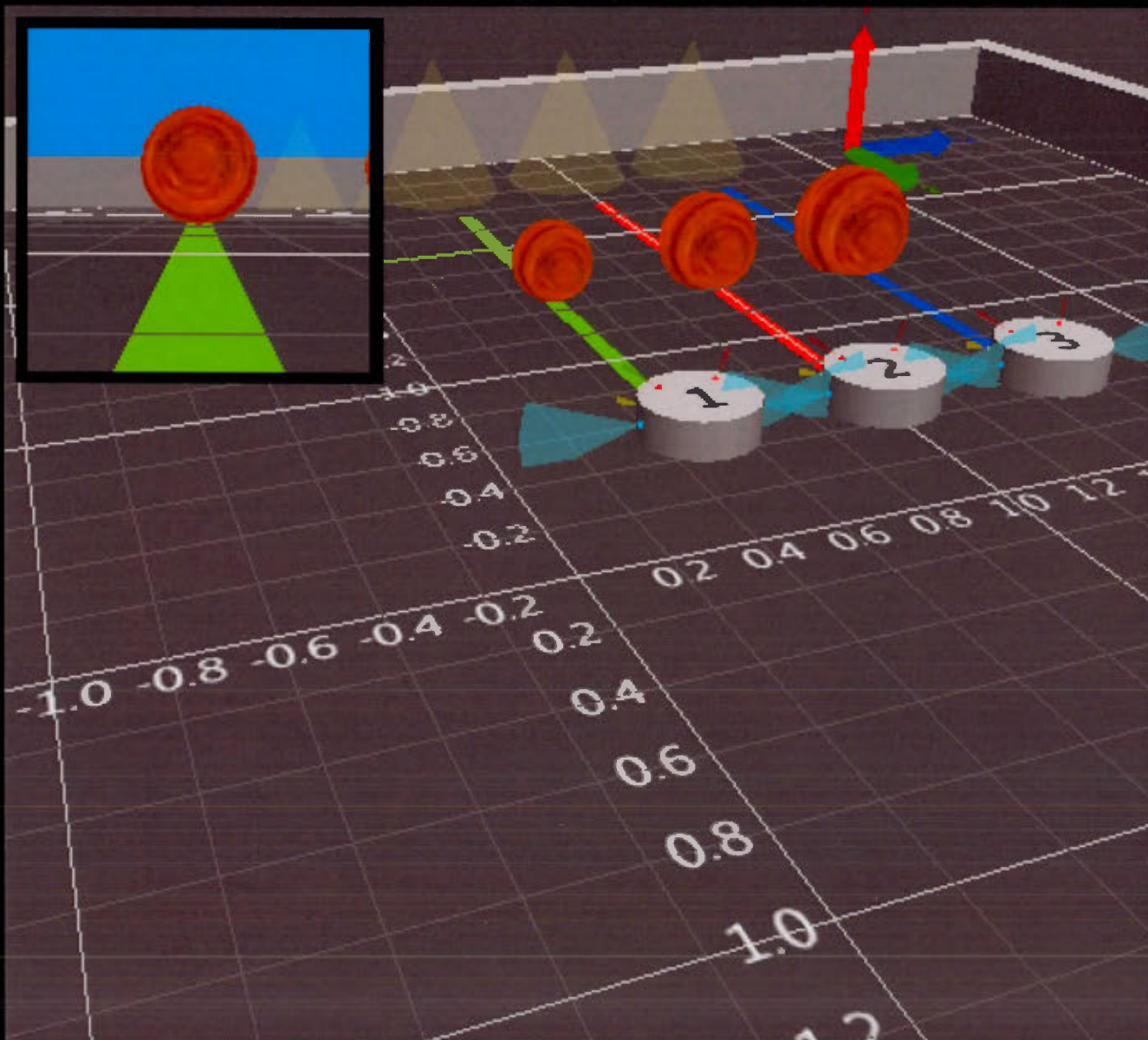
NeuroSim is not meant for exploration that focuses on particular low level details or going too deep in the comprehension of a given intelligent behaviour, and it may provide incoherent results if an attempt is made in these directions. As such, neural spikes and synapses represent convenient description models of input/output integration and memory traces, without requiring the simulation of the precise dynamical 3D processes of each protein for example. On the other hand, it may not be pertinent to simulate fined grained analog responses from a particular electronic robotic sensor component. Thus, a proper balance of the modeled object as well as the coherence of whole agent/environment is needed, focusing on concrete observable results of adaptive behavior of virtual robots, which is a premise to transfer the produced controllers in physical platforms.

In conclusion, we presented in this paper descriptions and issues regarding NeuroSim, a virtual 3D-world simulator and editor, dedicated to the investigation of the intelligence phenomenon through AI agents. This software module, part of the SIMCOG suite, bring novelty in the domain of bio-inspired robotic by merging ASNN controllers with current robotic components, exploring adaptive behaviors through the integration of multiple sensory modalities in a virtual 3D world.

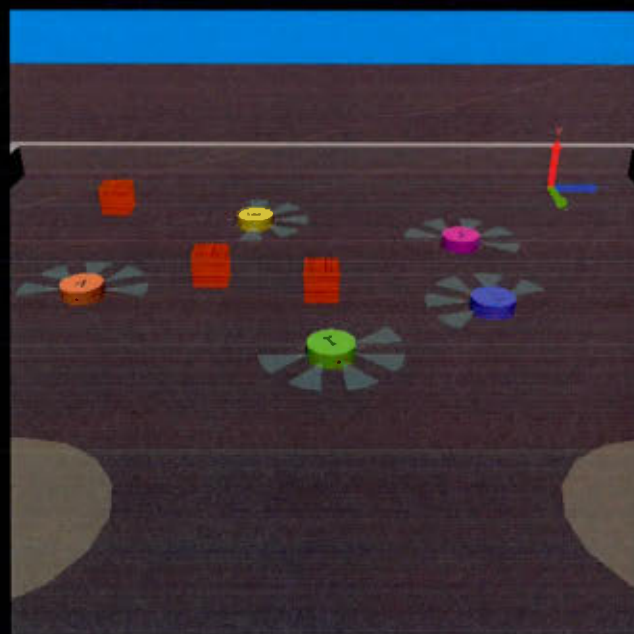
## References

- Alnajjar, F., Murase, K., 2008. A Simple Aplysia-Like Spiking Neural Network to generate Adaptive Behavior in Autonomous Robots, *Adaptive Behavior*, 16 (5), 306-324.
- Cyr, A., Boukadoum, M., Poirier, P., 2009. AI-SIMCOG : a simulator for spiking neurons and multiple animat's behaviours, *Neural Computation & Application*, 18 (5), 431-446.
- Cyr, A., Boukadoum, M., 2012. Classical conditioning in different temporal constraints: an STDP learning rule for robots controlled by spiking neural networks, *Adaptive Behavior*, 20 (4), 257-272.
- Floreano, D., Mattiussi, C., 2008. *Bio-Inspired artificial intelligence: theories, methods and technologies*, The MIT Press, Cambridge, MA.
- Gerstner, W., Kistler, W., 2002. *Spiking Neuron Models: single neurons, populations, plasticity*, Wulfram and Werner, Cambridge University Press.
- Hopkins, J.K., Spranklin, B.W., Gupta, S.K., 2009. A survey of snake-inspired robot designs, *Bioinspiration and Biomimetics*, 4 (2), 021001.
- Maass, W., Bishop, C.M., (eds), 1999. *Pulsed neural networks*, The MIT Press, Cambridge, MA.
- Mataric, M., 1997. Studying the role of embodiment in cognition, *Cybernetics and systems*, 28, 457-70.

- Paolo, E.A.D., 2003. Evolving spike-timing-dependent plasticity for single-trial learning in robots, *Philosophical Transactions of the Royal Society of London*, 361, 2299-2319.
- Pfeifer, R., Scheier, C., 1999. *Understanding Intelligence*, The MIT Press, Cambridge, MA.
- Soula, H., Alwan, A., Beslon, G., 2005. Learning at the edge of chaos : Temporal coupling of spiking neurons controller for autonomous robotic, *AAAI Spring symposium on developmental robotics*, Stanford, CA.
- Webb, B., Reeve, R., Horchler, A., Quinn, R., 2003. Testing a model of cricket phonotaxis on an outdoor robot platform, *4th British Conference on (Mobile) Robotics - Towards Intelligent Mobile Robots*.



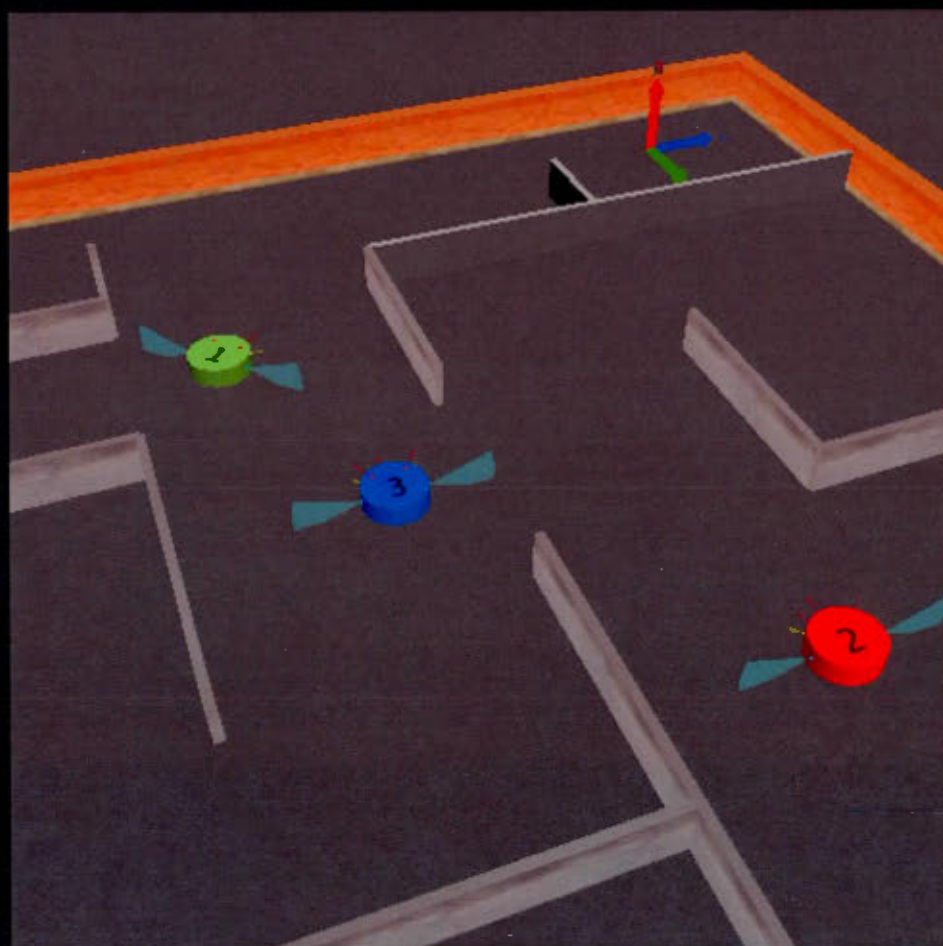
**Plate 1.**  
Screenshots of NeuroSim showing different  
3D views of the Virtual World during experiments.







**Plate 2.**  
Screenshots of NeuroSim showing  
different 3D views of the Virtual World  
during experiments.







## BIBLIOGRAPHIE

- Abbott, L.F. et Nelson, S.B. (2000). Synaptic plasticity: taming the beast. *Nature Neuroscience*, 3, 1178-1183.
- Abbott, L.F. et Regehr, W.G. (2004). Synaptic computation. *Nature*, 431, 796-803.
- Abeles, M. (1994). Firing rates and well-timed events. Dans: E. Domany, K. Schulten, and J.L. van Hemmen, Dir., *Models of neural networks 2*, chap. 3, 121-144, Springer, New York.
- Agin, V., Chichery, R., Dickel, L. et Chichery, M.-P. (2006). The “prawn-in-the-tube” procedure in the cuttlefish: habituation or passive avoidance learning? *Learn. Mem.*, 13, 97-101.
- Alexander, J., Audesirk, T.E. et Audesirk, G.J. (1984). One-trial reward learning in the snail *Lymnaea stagnalis*. *J. Neurobiol.*, 15, 67-72.
- Alexander, W.H. (2007). Shifting attention using a temporal difference prediction error and high-dimensional input. *Adaptive Behavior*, 15(2), 121-133.
- Alnajjar, F. et Murase, K. (2006). Self-organization of spiking neural network that generates autonomous behavior in a real mobile robot. *International journal of neural systems*, 16(4), 229-239.
- Alnajjar, F. et Murase, K. (2008). A simple Aplysia-Like spiking neural network to generate adaptive behavior in autonomous robots. *Adaptive Behavior*, 16(5), 306-324.
- Alnajjar, F., Zin, I.B.M. et Murase, K. (2009). A Hierarchical Autonomous Robot Controller for Learning and Memory: Adaptation in a Dynamic Environment. *Adaptive Behavior*, 17(3), 179-196.
- Alonso, L., Moreno, R., Vázquez, M. et Santacreu, J. (2005). Simulation of the filtering role of habituation to stimuli. *The Spanish journal of psychology*, 8(2), 134-141.
- Anderson, M.L. (2003). Embodied cognition: A field guide. *Artificial Intelligence*, 149, 91-130.
- Andrew, R.J. et Savage, H. (2000). Appetitive learning using visual conditioned stimuli in the pond snail, *Lymnaea*. *Learn. Mem.*, 73, 258-273.
- Antonov, I., Antonova, I., Kandel, E.R. et Hawkins, R.D. (2001). The contribution of activity-dependent synaptic plasticity to classical conditioning in Aplysia. *J. Neurosci.*, 21(16), 6413-6422.
- Antonov, I., Kandel, E.R. et Hawkins, R.D. (2010). Presynaptic and postsynaptic mechanisms of synaptic plasticity and metaplasticity during intermediate-term memory formation in Aplysia. *J. Neurosci.*, 30(16), 5781-5791.
- Ardiel, E.L. et Rankin, C.H. (2010). An elegant mind: Learning and memory in *Caenorhabditis elegans*. *Learn. Mem.*, 17, 191-201.
- Aréna, P., Fortuna, L., Frasca, M. et Patané, L. (2009). Learning anticipation via spiking networks: application to navigation control. *IEEE Trans. Neural. Netw.*, 20(2), 202-216.
- Bailey, C. H., Bartsch, D. et Kandel, E. R. (1996). Toward a molecular definition of long-term memory storage. *Proceedings of the National Academy of Sciences, U.S.A.*, 93, 13445-13452.

- Balkenius, C. et Morén, J. (1998). Computational models of classical conditioning: a comparative study. Dans: Mayer, J.-A., Roitblat, H. L., Wilson, S. W., and Blumberg, B. (Dir.), *From Animals to Animats 5*. Cambridge, MA: MIT Press.
- Balkenius, C. (2000). Attention, Habituation and conditioning: toward a computational model. *Cognitive Science Quarterly*, 1(2), 171-204.
- Baxter, D.A. et Byrne, J.H. (2006). Feeding behavior of Aplysia: A model system for comparing cellular mechanisms of classical and operant conditioning. *Learn. Mem.*, 13, 669-680.
- Baxter, D.A., Cataldo, E. et Byrne, J.H. (2013). Computational analyses of learning networks. Dans: *Invertebrate Learning and Memory*, Dir., Menzel, R. et Benjamin, P.R., Elsevier Limited., Ch. 7, pp. 69-78.
- Beck, C.D.O. et Rankin, C.H. (1995). Heat-shock disrupts long-term memory consolidation in *Caenorhabditis elegans*. *Learn. Mem.*, 2, 161-177.
- Beck, C.D.O. et Rankin, C.H. (1997). Long-term habituation is produced by distributed training at long ISIs and not by massed training or short ISIs in *Caenorhabditis elegans*. *Anim. Learn. Behav.*, 25, 446-457.
- Bédécarrats, A., Cornet, C., Simmers, J. et Nargeot, R. (2013). Implication of dopaminergic modulation in operant reward learning and the induction of compulsive-like feeding behavior in Aplysia. *Learn. Mem.*, 20, 318-327.
- Beer, R.D. (1995). A dynamical systems perspective on agent-environment interaction. *Artificial intelligence*, 72, 173-215.
- Bekey, G.A. (2005). *Autonomous robots: From biological inspiration to implementation and control*. Cambridge, MA: MIT Press.
- Bell, C.C., Han, V.Z., Sugawara, Y. et Grant, K. (1997). Synaptic plasticity in a cerebellum-like structure depends on temporal order. *Nature*, 387, 278-281.
- Benda, J. (2002). Dynamics of spike-frequency adaptation. Dans: *Single neuron dynamics: models linking theory and experiment*. Chap.4, dissertation. Berlin, Germany.
- Bi, G. Q. et Poo, M. M. (1998). Synaptic modifications in cultured Hippocampal neurons: dependence on spike timing, synaptic strength, and postsynaptic cell type. *J. Neurosci.*, 18, 10464-72.
- Bi, G. Q. et Poo, M. M. (2001). Synaptic modification by correlated activity: Hebb's postulate revisited. *Annu. Rev. Neurosci.*, 24, 139-66.
- Bialek, W. et Zee, A. (1990). Coding and computation with neural spike trains. *Journal of statistical physics*, 59, 103-115.
- Bienenstock, E.L., Cooper, L.N. et Munro, P.W. (1982). Theory for the development of neuron selectivity: Orientation specificity and binocular interaction in visual cortex. *J. Neurosci.*, 2, 32-48.
- Bliss, T.V.P. et Lømo, T. (1973) Long-lasting potentiation of synaptic transmission in the dentate area of the anaesthetized rabbit following stimulation of the perforant path. *J. Physiol. (Lond.)*, 232, 331-356.
- Bliss, T.V.P. et Collingridge, G.L. (1993). A synaptic model of memory: long-term potentiation in the hippocampus. *Nature*, 361, 31-39.
- Bohte, S.M. (2004). The evidence for neural information processing with precise spike-times: A survey. *Natural computing*, 3(2), 195-206.

- Bohte, S.M. et Mozer, M.C. (2007). Reducing de variability of neural responses: A computational theory of spike-timing dependant plasticity. *Neural Computation*, 19, 371-403.
- Bongard, J.C. (2013). Evolutionary robotics. *Commun. ACM* 56, 74-85. doi: 10.1145/2492007.2493883.
- Bozorgmehr, T., Ardiel, E.L., McEwan, A.H. et Rankin, C.H. (2013). Mechanisms of plasticity in a *Caenorhabditis elegans* mechanosensory circuit. *Front. Physiol.*, (4)88, 1-11. doi:10.3389/fphys.2013.00088.
- Brembs, B., Lorenzetti, F.D., Reyes, F.D., Baxter, D.A. et Byrne, J.H. (2002). Operant reward learning in *Aplysia*: Neuronal correlates and mechanisms. *Science*, 296, 1706-1709.
- Brembs, B. (2003). Operant conditioning in invertebrates. *Curr. Opin. Neurobiol.*, 13, 710-717.
- Brembs, B. et Plendl, W. (2008). Double dissociation of PKC and AC manipulations on operant and classical learning in *drosophila*. *Current Biology*, 18, 1168-1171.
- Brette, R. et Guigon, E. (2003). Reliability of spike timing is a general property of spiking model neurons. *Neural computation*, 15(2), 279-308.
- Brette, R. (2004). Dynamics of one-dimensional spiking neuron models. *J. Math. Biol.* 48(1), 38-56.
- Bristol, A.S. et Carew, T.J. (2005). Differential role of inhibition in habituation of two independent afferent pathways to a common motor output. *Learn. Mem.*, 12, 52-60.
- Brooks, R.A. (1986). A robust layered control system for a mobile robot. *IEEE Journal of Robotics and Automation*, 2, 14-23.
- Brooks, R.A. (1999). *Cambrian intelligence*. Cambridge, MA: MIT Press.
- Brooks, R.A. (1991a). Intelligence without reason. Dans: *Proceedings of the 12<sup>th</sup> Intl. Conference on artificial intelligence (IJCAI-91)*, Morgan Kaufmann, San Mateo, CA, 569-595.
- Brooks, R.A. (1991b). Intelligence without representation. *Artificial Intelligence*, 47, 139-159.
- Broster, B.S. et Rankin, C.H. (1994). Effects of changing interstimulus interval during habituation in *Caenorhabditis elegans*. *Behav. Neurosci.*, 108(6), 1019-1029.
- Bryson, A.E. and Y. Ho, Y. (1969). *Applied Optimal Control*, Blaisdell Publishing Company, Waltham, Massachusetts.
- Bryson, J.J. (2007). Mechanisms of action selection: Introduction to the special issue. *Adaptive Behavior*, 15(1), 5-8.
- Bristol, A.S. et Carew, T.J. (2005). Differential role of inhibition in habituation of two independent afferent pathways to a common motor output. *Learn. Mem.* 12, 52-60.
- Byrne, J.H. (1982). Analysis of synaptic depression contributing to habituation of gill-withdrawal reflex in *Aplysia californica*. *J. Neurophysiol.*, 48, 431-438.
- Cachón, A. et Vázquez, R.A. (2015). Tuning the parameters of an integrate and fire neuron via a genetic algorithm for solving pattern recognition problems. *Neurocomputing*, 148, 187-197.
- Cai, D., Pearce, K., Chen, S. et Glanzman, D.L. (2012). Reconsolidation of long-term memory in *Aplysia*. *Curr. Biol.*, 22, 1783-1788.
- Cai, S.Q., Wang, Y., Park, K.H., Tong, X., Pan, Z. et Sesti, F. (2009). Auto-phosphorylation of a voltage-gated K<sup>+</sup> channel controls non-associative learning. *EMBO J.*, 28(11), 1601-1611.



- Caporale, N. et Dan, Y. (2008). Spike timing-dependent plasticity: A Hebbian learning rule. *Annual review of neuroscience*, 31(1), 25-46.
- Carlson, K.D., Nageswaran, J.M., Dutt, N. et Krichmar, J.L. (2014). An efficient automated parameter tuning framework for spiking neural networks. *Frontiers in neuroscience*, doi: 10.3389/fnins.2014.00010.
- Castelluci, V.F. et Kandel, E.R. (1974). A quantal analysis of the synaptic depression underlying habituation of the gill-withdrawal reflex in *Aplysia*. *Proc. Natl. Acad. Sci. USA*, 71(12), 5004-5008.
- Çevik, M.O. (2014). Habituation, sensitization, and Pavlovian conditioning. *Frontiers in integrative neuroscience*, 8(13), 1-6.
- Chalasani, S.H., Chronis, N., Tsunozaki, M., Gray, J.M., Ramot, D., Goodman, M.B. et Bargmann, C.I. (2007). Dissecting a circuit for olfactory behaviour in *Caenorhabditis elegans*. *Nature*, 450, 63-70, doi:10.1038/nature06292.
- Chalfie, M., Sulston, J.E., White, J.G., Southgate, E., Thomson, J.N. et Brenner, S. (1985). The neural circuit for touch sensitivity in *Caenorhabditis elegans*. *Journal of Neurosciences*, 5(4), 956-964.
- Chandana, P. (2004). *Morphology and computation*, Dans: Proceedings of the eighth international conference on the simulation of adaptive behavior, From animals to animats, Dir: Stefan Schaal, Auke Ijspeert, Aude Billard, Sethu Vijayakumar, John Hallam, and Jean-Arcady Meyer, MIT Press, 33-38.
- Chang, C. (2005). Using sensor habituation in mobile robots to reduce oscillatory movements in narrow corridors. *IEEE Trans. Neural Netw.*, 16(6), 1582-1589.
- Chartier, S. et Boukadoum M. (2006). A Sequential Dynamic Heteroassociative Memory for Multistep Pattern Recognition and One-to-Many Association. *IEEE Trans. neural netw.*, 17(1), 59-68.
- Chiel, H. J. et Beer, R. D. (1997). The brain has a body: Adaptive behavior emerges from interactions of nervous system, body and environment. *Trends in Neurosciences*, 20(12), 553-557.
- Chrisley, R. (2003). Embodied artificial intelligence. *Artificial Intelligence*, 149(1), 131-150.
- Christodoulou, C., Bugmann, G. et Clarkson, T.G. (2002). A spiking neuron model: Applications and learning. *Neural Netw.*, 15, 891-908.
- Clark, A. (1997). *Being There: Putting brain, body and world together again*. Cambridge, MA: MIT Press.
- Crevier, D. (1997). *À la recherche de l'intelligence artificielle*. France, Flammarion.
- Dan, Y. et Poo, M.-M. (2004). Spike timing-dependent plasticity of neural circuits. *Neuron*, 44, 23-30.
- Dan, Y. et Poo, M.-M. (2006). Spike timing-dependant plasticity: From synapse to perception. *Physiological Reviews*, 86, 1033-1048.
- Davis, M. (1970). Effects on interstimulus interval length and variability on startle-response habituation in the rat. *J. Comp. Physiol. Psychol.*, 72(2), 177-192.
- Déniz, O., Lorenzo, J. et Hernández, M. (2004). A simple habituation mechanism for perceptual user interfaces. *Inteligencia Artificial*, Revista Iberoamericana de inteligencia artificial, 23, 87-96.
- Di Paolo, E.A. (2002). Spike timing dependent plasticity for evolved robots. *Adaptive Behavior*, 10(3-4), 243-263.

- Di Paolo, E.A. (2003). Evolving spike-timing-dependent plasticity for single-trial learning in robots. *Philosophical Transactions of the Royal Society of London A*, 361, 2299-2319.
- Doncieux, S., Bredeche, N., Mouret, J.-B. et Eiben, A.E. (2015). Evolutionary robotics : what, why and where to. *Frontiers in robotics and AI*. doi:10.3389/frobt.2015.00005.
- Dong, S. et Clayton, D.F. (2009). Habituation in songbirds. *Neurobiol. Learn. Mem.*, 92(2), 183-188.
- Dreyfus, G., Martinez, J.-M., Samuelides, M., Gordon, M.B., Badran, F., Thiria, S. et Hérault, L. (2002). *Réseaux de neurones : méthodologie et applications*. Éditions Eyrolles, Paris.
- Engel, J.E. et Wu, C.F. (2009). Neurogenetic approaches to habituation and dishabituation in *Drosophila*. *Neurobiol. Learn. Mem.*, 92(2), 166-175.
- Ezzeddine, Y. et Glanzman, D.L. (2003). Prolonged habituation of the gill-withdrawal reflex in *Aplysia* depends on protein synthesis, protein phosphatase activity, and postsynaptic glutamate receptors. *Journal of Neuroscience*, 23(29), 9585-9594.
- Faumont, S., Lindsay, T.H. et Lockery, S.R. (2012). Neuronal microcircuits for decision making in *C. elegans*. *Current opinion in neurobiology*, 22(4), 580-591.
- Feldman, D.E. (2012). The spike-timing dependence of plasticity. *Neuron*, 75(4), 556-571.
- Fiore, V.G., Mannella, F., Mirolli, M., Gurney, K. et Baldassarre, G. *Instrumental conditioning driven by neutral stimuli: A model tested with a simulated robotic rat*. Dans: Schlesinger, M., Berthouze, L., and Balkenius, C. (2008) Proceedings of the Eighth International Conference on Epigenetic Robotics: Modeling Cognitive Development in Robotic Systems. Lund University Cognitive Studies, 139.
- Fischer, T.M., Jacobson, D.A. et Demorest-Hayes, K. (2014). Network processes involved in the mediation of short-term habituation in *Aplysia*: Contribution of intrinsic regulation of excitability and synaptic augmentation. *Front. Integr. Neurosci.*, 15, 1-12. doi: 10.3389/fnint.2014.00015.
- Floreano, D. et Mattiussi, C. (2008). *Bio-Inspired artificial intelligence: Theories, methods and technologies*. Cambridge, MA: MIT Press.
- Fontana, A., Soltoggio, A. et Wrobel, B. (2014). POET: an evo-devo method to optimize the weights of a large artificial neural networks. Dans: *Proceedings of the Fourteenth International Conference on the Synthesis and Simulation of Living Systems (ALIFE XIV)*. Cambridge, MA: MIT Press.
- Freeman, W.J. (1994). Neural networks and chaos. *Journal of theoretical biology*, 171, 13-18.
- Frick, L.R., Bernadez-Vidal, M., Hocht, C., Zanutto, B.S. et Rapanelli, M. (2014). Dual role of serotonin in the acquisition and extinction of reward-driven learning: Involvement of 5-HT1A, 5-HT2A and 5-HT3 receptors. *Behav. Brain Res.*, 15(277), 193-203.
- Froemke, R.C. et Dan, Y. (2002). Spike-Timing-Dependent synaptic modification induced by natural spike trains. *Nature*, 416, 433-438.
- Froemke, R., Poo, M.M. et Dan, Y. (2005). Spike-timing-dependent synaptic plasticity depends on dendritic location. *Nature*, 434, 221-225.
- Frost, L., Kaplan, S.W., Cohen, T.E., Henzi, V., Kandel, E.R. et Hawkins, R.D. (1997). A simplified preparation for relating cellular events to behavior: contribution of LE and unidentified siphon sensory neurons to mediation and habituation of the *Aplysia* gill-and siphon-withdrawal reflex. *J. Neurosci.*, 17(8), 2900-2913.
- Gaudiano, P. et Chang, C. (1997). *Adaptive obstacle avoidance with a neural network for operant conditioning: Experiments with real robots*. Dans: IEEE, Proceedings of CIRA'97, USA.

- Genesis [s.d.], *The GENESIS 2 simulator home page*. Récupéré de <http://genesis-sim.org>
- Gerstner, W., Kempter R., Van Hemmen, J.L. et Wagner, H. (1996). A neuronal learning rule for sub-millisecond temporal coding. *Nature*, 386, 76-78.
- Gerstner, W. et Kistler, W. (2002). *Spiking Neuron Models: Single neurons, populations, plasticity*. Wulfram and Werner, Cambridge University Press.
- Giles, A.C. et Rankin, C.H. (2009). Behavioral and genetic characterization of habituation using *Caenorhabditis elegans*. *Neurobiol. Learn. Mem.*, 92, 139-146.
- Gjorgjieva, J., Clopath, C., Audet, J. et Pfister, J.P. (2011). A triplet spike-timing-dependent plasticity model generalizes the Bienenstock–Cooper–Munro rule to higher-order spatiotemporal correlations. *Proc. Natl. Acad. Sci. USA*, 108(48), 19383–19388.
- Glanzman, D.L. (2010). Common mechanisms of synaptic plasticity: Minireview in vertebrates and invertebrates. *Curr. Biol.*, 20, 31-36.
- Grillner, S. (1985) Neurobiological bases of rhythmic motor acts in vertebrates. *Science*, 228, 143-149.
- Grossberg, S. (1982). *Studies of mind and brain: neural principles of learning, perception, development, cognition and motor control*. Boston Studies in the Philosophy of Science, 70, Dordrecht, Holland: Reidel.
- Groves, P.M. et Thompson, R.F. (1970). Habituation: A dual-process theory. *Psychol. Rev.*, 77(5), 419-450.
- Hawkins, R.D., Kandel, E.R. et Bailey, C.H. (2006). Molecular mechanisms of memory storage in *Aplysia*. *Biol. Bull.*, 210, 174-191.
- Hebb, D.O. (1949). *The organization of behavior*. Wiley, New York.
- Hills, T., Brockie, P.J. et Maricq, A.V. (2004). Dopamine and glutamate control area-restricted search behavior in *Caenorhabditis elegans*. *J. Neurosci.*, 24(5), 1217-1225.
- Hodgkin, A.L. et Huxley, A.F. (1952). A quantitative description of ion currents and its applications to conduction and excitation in nerve membranes. *Journal of Physiology(Lond.)*, 117, 500-544.
- Holland, O. (2003). The first biologically inspired robots. *Robotica*, 21(4), 351–363.
- Hopfield, J.J. (1982). Neural networks and physical systems with emergent collective computational abilities. *Proceedings of the national academy of the sciences of the USA*, 79, 2554-2558.
- Hull, C.L. (1932). The goal-gradient hypothesis and maze learning. *Psychol. Rev.*, 39(1), 25-43.
- Humphries, M.D., Gurney, K. et Prescott, T.J. (2007). Is there a brainstem substrate for action selection? *Philosophical Transactions of the Royal Society B: Biological Sciences*, 362(1485), 1627-1639.
- Innis, N.K. et Staddon, J.E.R. (1989). What should comparative psychology compare? *International Journal of Comparative Psychology*, 2(3), 145-156.
- Izhikevich, E.M. (2003). Simple model of spiking neurons. *IEEE Transactions on Neural Networks*, 14, 1569-1572.
- Izhikevich, E.M. (2004). Which model to use for cortical spiking neurons? *IEEE Trans. Neural. Netw.*, 15(5), 1063-1070.
- Izhikevich, E.M. (2007). Solving the distal reward problem through linkage of STDP and dopamine signaling. *Cerebral Cortex*, 17(10), 2443-2452.
- Jacob, V., Brasier, J., Erchova, I., Feldman, D. et Schulz, D.E. (2007). Spike timing-dependent synaptic depression in the in vivo barrel cortex of the rat. *J. Neurosci.*, 27(6), 1271-84.



- JMonkeyEngine [s.d.], *A cross-platform game engine for adventurous Java developers*. Récupéré de <http://jmonkeyengine.org>
- Jones, N.G., Kemenes, I., Kemenes G. et Benjamin, P.R. (2003) A persistent cellular change in a single modulatory neuron contributes to associative long-term memory. *Curr. Biol.*, 13(12), 1064-1069.
- Kandel, E.R. (2001). The molecular biology of memory storage: a dialogue between genes and synapses. *Science*, 294(5544), 1030-1038.
- Kandel, E.R., Schwartz, J.H., Jessell, T.M., Siegelbaum, S. et Hudspeth, A.J. (Dir.) (2012). *Principles of neural science, Fifth edition*. McGraw-Hill: NY, USA.
- Kempler, R. Gerstner, W., van Hemmen, J.L. et Wagner, H. (1996). Temporal coding in the sub-millisecond range: Model of barn owl auditory pathway. Dans: *Advances in Neural Information Processing Systems 8*, 124-130, NIPS, Cambridge, MA, MIT Press.
- Kier, R.J., Ames, J.C., Beer, R.D. et Harrison, R.R. (2006). Implementation of multipattern generators in analog VLSI. *IEEE Transactions on neural networks*, 17(4), 1025-1038.
- Kindt, K.S., Quast, K.B., Giles, A.C., Hendrey, D., Nicastro, I., Rankin, C.H. et Schafer, W.R. (2007). Dopamine mediates context-dependent modulation of sensory plasticity in *C. elegans*. *Neuron*, 55(4), 662-676.
- Kitamura, K.I., Amano. S. et Hosono, R. (2001). Contribution of neurons to habituation to mechanical stimulation in *Caenorhabditis elegans*. *J. Neurobiol.*, 46(1), 29-40.
- Koch, C. (1999) *Biophysics of computation: information processing in single neurons*. Oxford university press, New York.
- Kohonen, T. (1982). Self-organized formation of topologically correct feature maps. *Biological Cybernetics*, 43, 59-69.
- Krasne, F.B. et Teshiba, T.M. (1995). Habituation of an invertebrate escape reflex due to modulation by higher centers rather than local events. *Proc. Natl. Acad. Sci. USA*, 92, 255(5052), 3362-3366.
- Lakhtakia, A. et Martin-Palma, R. (Dir.) (2013). *Engineered Biomimicry*. MA: Elsevier.
- Lang, K.J. et Hinton, G.E. (1988). *The development of the time-delay neural network architecture for speech recognition*. Technical report CMU-CS-88-152, Carnegie-Mellon University, Pittsburgh, PA.
- Larkin, A., Karak, S., Priya, R., Das, A., Ayyub, C., Ito, K., Rodrigues, V. et Ramaswami, M. (2010). Central synaptic mechanisms underlie short-term olfactory habituation in *Drosophila* larvae. *Learn. Mem.*, 17(12), 645-653.
- Lisman, J.E. et Idiart, M.A. (1995). Storage of 7+/- 2 short-term memories in oscillatory subcycles. *Science*, 267, 1512-1515.
- Liu, C. et Shapiro, J. (2007). *Implementing Classical Conditioning with Spiking Neurons*. Dir. J. Marques de Sa et al, ICANN 2007, Part I, LNCS, 4668, pp.400-410. Springer-Verlag, Berlin.
- Liu, Y. et Sun, D. (Dir.) (2012). *Biologically inspired robotics*. NW: CRC Press.
- Lloyd, D.R., Medina, D.J., Hawk, L.W., Fosco, W.D. et Richards, J.B. (2014). Habituation of reinforcer effectiveness. *Front. Integr. Neurosci.*, 7, 1-20, doi: 10.3389/fnint.2013.00107.
- Lorenzetti, F.D., Baxter, D.A. et Byrne, J.H. (2008). Molecular mechanisms underlying a cellular analog of operant reward learning. *Neuron*, 59(5), 815-828.
- Ludobots [s.d.], *Education in Evolutionary Robotics*. Récupéré de <http://www.uvm.edu/~ludobots>



- Luger, G.F. (2002). *Artificial intelligence: Structures and strategies for complex problem solving* (4e édition). Edinburgh: Gate, Harlow: Pearson Education Limited.
- Lüscher C. et Malenka R.C. (2012). NMDA receptor-dependent long-term potentiation and long-term depression (LTP/LTD). *Cold Spring Harb Perspect Biol* doi: 10.1101/cshperspect.a005710.
- Maass, W. (1997). Networks of spiking neurons: the third generation of neural network models. *Neural Networks*, 10(9), 1659-1671.
- Maass, W. et Bishop, C.M. (Dir.) (1999). *Pulsed neural networks*. Cambridge, MA: Bradford Book, MIT Press.
- Markram, H., Lübke, J., Frotscher, M. et Sakmann, B. (1997). Regulation of synaptic efficacy by coincidence of postsynaptic APs and EPSEs. *Science*, 275(5297), 213-215.
- Markram, H., Gerstner, W. et Sjöström, P.J. (2011). A history of spike-timing-dependent plasticity. *Frontiers in synaptic neuroscience*, 3(4), 1-24.
- Marsland, S., Nehmzow, U. et Shapiro, J. (1999). A model of habituation applied to mobile robots. Dans: *Proceedings of TIMR 99*, Towards Intelligent Mobile Robots, Bristol.
- Marsland, S., Nehmzow, U. et Shapiro, J. (2005). On-line novelty detection for autonomous mobile robots. *Robotics and autonomous systems*, 51(2), 191-206.
- Marsland, S. (2009). Using habituation in machine learning. *Neurobiol. Learn. Mem.*, 92(2), 260-266.
- Masini, C.V., Day, H.E.W. et Campeau, S. (2008). Long-term habituation to repeated loud noise is impaired by relatively short interstressor intervals in rats. *Behav Neurosci.*, 122(1), 210-223.
- Mataric, M. (1997). Studying the role of embodiment in cognition. *Cybernetics and systems* 28, 457-70.
- MathWorks [s.d.], *MATLAB Speaks Machine Learning*. Récupéré de <http://mathworks.com>
- Matlin, M.W. (2005). *Cognition, sixth edition*. Danvers: MA: John Wiley & Sons, Inc.
- Mayford, M., Siegelbaum, S.A. et Kandel, E.R. (2012). Synapses and memory storage. *Cold Spring Harb Perspect Biol* 4(6):a005751.
- Mazur, J.E. (2013). *Learning and behavior*, 7<sup>th</sup> ed., Pearson Education Inc., NJ: USA.
- McCulloch, W.S. et Pitts, W.H. (1943). A logical calculus of the ideas immanent in nervous activity. *Bulletin of mathematical biophysics*, 5, 115-133.
- MCell [s.d.], *Monte Carlo Cell*. Récupéré de <http://mcell.org>
- McSweeney, F.K. et Murphy, E.S. (2009). Sensitization and habituation regulate reinforcer effectiveness. *Neurobiol. Learn. Mem.*, 92(2), 189-198. doi: 10.1016/j.nlm.2008.07.002.
- McNamara, A.M., Magidson, P.D., Linster, C., Wilson, D.A. et Cleland, T.A. (2008). Distinct neural mechanisms mediate olfactory memory formation at different timescales. *Learn. Mem.*, 15(3), 117-125.
- Menzel, R. et Benjamin, P.R. (Dir.) (2013). *Invertebrate learning and memory*. Handbooks of behavioral neuroscience. Academic Press Elsevier, San Diego, USA.
- Mehrtash, N., Jung, D., Hellmitch, H.H., Schoenauer, T., Lu, V. et Klar, H. (2003). Synaptic plasticity in spiking neural networks: A system approach. *IEEE Trans. Neural Netw.*, 14(5), 980-992.
- Michel, M., Kemenes, I., Müller, U. et Kemenes, G. (2008). Different phases of long-term memory require distinct temporal patterns of PKA activity after single-trial classical conditioning. *Learn. Mem.*, 15, 694-702.

- Microsoft Robotics [s.d.], *Microsoft Robotics Developer Studio 4*. Récupéré de <http://www.microsoft.com/robotics>
- Minsky, M. et Papert, S. (1969). *Perceptrons*. Cambridge, MA: MIT Press.
- Morris, R.G., Anderson, E., Lynch, G.S. et Baudry, M. (1986). Selective impairment of learning and blockade of long-term potentiation by an N-methyl-D-aspartate receptor antagonist, AP5. *Nature*(6056), 319, 774–776.
- Mozzachiodi, R., Baxter, D.A. et Byrne, J.H. (2013). Comparison of operant and classical conditioning in the feeding system of *Aplysia*. Dans: *Invertebrate Learning and Memory*, Dir. R. Menzel et P. Benjamin, Elsevier Limited.
- Natschläger, T., Maass, W. et Zador, A. (2001). Efficient temporal processing with biologically realistic dynamic synapses. *Network: computation in neural systems* 12, 75-87.
- Naundorf, B., Wolf, B. et Volgushev, M. (2006). Unique features of action potential initiation in cortical neurons. *Nature*, 440, 1060-1063.
- Neuron [s.d.], *NEURON for empirically-based simulations of neurons and networks of neurons*. Récupéré de <http://www.neuron.yale.edu/neuron>
- Nolfi, S. et Floreano D. (2000). *Evolutionary Robotics: The biology, intelligence and technology of self-organizing machines*. Cambridge, MA: MIT Press.
- Nuttley, W.M., Harbinder, S. et van der Kooy, D. (2001). Regulation of distinct attractive and aversive mechanisms mediating benzaldehyde chemotaxis in *Caenorhabditis elegans*. *Learn. Mem.*, 8(3), 170-181.
- Ode4j [s.d.], *Java 3D Physics Engine & Library*. Récupéré de <http://www.ode4j.org>
- O'Keefe, J.D.J. (1971). The hippocampus as a spatial map. Preliminary evidence from unit activity in the freely-moving rat. *Brain Res.*, 34(1), 171–175.
- O'Keefe, J. et Recce, M.L. (1993). Phase relationship between hippocampal place units and the EEG theta rhythm. *Hippocampus*, 3, 317-30.
- OpenSimulator [s.d.], *What is OpenSimulator?*. Récupéré de <http://opensimulator.org>
- Pan, B. et Zucker, R.S. (2009). A general model of synaptic transmission and short-term plasticity. *Neuron*, 62(4), 539-554.
- Päpper, M., Kempter, R. et Leibold, C. (2011). Synaptic tagging, evaluation of memories, and the distal reward problem. *Learn. Mem.*, 18(1), 58-70.
- Pavlov, I. P. (1927). *Conditioned Reflexes: An Investigation of the Physiological Activity of the Cerebral Cortex*. Dir. G. V. Anrep. London, Oxford University Press.
- Pfeifer, R. et Scheier, C. (1999). *Understanding intelligence*. Cambridge, MA: MIT Press.
- Pfeifer, R. (2007). *How the body shapes the way we think: A new view of intelligence*. Cambridge, MA: MIT Press.
- Pfister, J.-P. et Gerstner, W. (2006). Triplets of spikes in a model of spike timing-dependent plasticity. *J. Neurosci.*, 26(38), 9673-9682.
- Pilz, P.K.D., Arnold, S.W., Rischawy, A.T. et Plappert, C.F. (2014). Long-term habituation of the startle response in mice is stimulus modality, but not context specific. *Front. Integr. Neurosci.*, 7, 1-10, doi: 10.3389/fnint.2013.00103.
- Pinsker, H., Kupfermann, I., Castellucci, V. et Kandel, E.R. (1970). Habituation and dishabituation in the gill withdrawal reflex in *Aplysia*. *Science*, 167(926), 1740–1742.
- Ponulak, F. et Kasinski, A. (2011). Introduction to spiking neural networks: Information processing, learning and applications. *Acta Neurobiol Exp*, 71, 409-433.
- Poznanski, R.R. (Dir.) (1999). *Modeling in the neurosciences: from ionic channels to neural networks*. India: Hardwood academic publishers.

- Qin, J. et Wheeler, A.R. (2007). Maze exploration and learning in *C. elegans*. *Lab Chip*, 7, 186-192.
- Rabinovitch, I., Chatzigeorgiou, M., Zhao, B., Treinin, M. et Shafer, W.R. (2014). Rewiring neural circuits by the insertion of ectopic electrical synapses in transgenic *C. elegans*. *Nat Commun.*, 5, 4442.
- Rankin, C.H. et Carew, T.J. (1987). Development of learning and memory in *Aplysia*. II. Habituation and dishabituation. *The journal of neuroscience*, 7(1), 133-143.
- Rankin, C.H. et Carew, T.J. (1988). Dishabituation and sensitization emerge as separate processes during development in *Aplysia*. *Journal of neuroscience*, 8(1), 197-211.
- Rankin, C.H., Beck, C.D.O. et Chiba, C.M. (1990). *Caenorhabditis elegans*: A new model system for the study of learning and memory. *Behav. Brain Res.*, 37(1), 89-92.
- Rankin, C.H. et Broster, B.S. (1992). Factors affecting habituation and recovery from habituation in the nematode *Caenorhabditis elegans*. *Behav. Neurosci.*, 106, 239-242.
- Rankin, C.H. (2000). Context conditioning, in habituation in the nematode *Caenorhabditis elegans*. *Behav. Neurosci.*, 114(3), 496-505.
- Rankin, C.H., Abrams, T., Barry, R.J., Bhatnagar, S., Clayton, D.F., Colombo, J., Coppola, G., Geyer, M.A., Glanzman, D.L., Marsland, S., McSweeney, F.K., Wilson, D.A., Wu, C.F. et Thompson, R.F. (2009). Habituation revisited: An updated and revised description of the behavioral characteristics of habituation. *Neurobiol. Learn. Mem.*, 92(2), 135-138.
- Redondo, R.L. et Morris, R.G. (2011). Making memories last: the synaptic tagging and capture hypothesis. *Nature Reviews*, 12(1), 17-30.
- Rescorla, R.A. et Wagner, A.R. (1972). *A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and non-reinforcement*. Dans: *Classical conditioning II: Current research and theory* (Dir.: A.H. Black, W.F. Prokasy), 64-99. Appleton-Century-Crofts, New York.
- Riegler, A. (2002). When is a cognitive system embodied? *Cognitive system research*, Special issue on "Situated and embodied cognition", 3, 339-348.
- Rieke, F., Warland, D., de Ruyter van Steveninck, R et Bialek, W. (1997). *Spikes: Exploring the Neural Code*. MIT Press, Cambridge, MA.
- Roberts, P. D. et Bell, C. C. (2002). Spike timing dependent synaptic plasticity in biological systems. *Biological Cybernetics*, 87, 392-403.
- Rose, J.K. et Rankin, C.H. (2001). Analyses of habituation in *Caenorhabditis elegans*. *Learn. Mem.*, 8(2), 63-69.
- Rose, J.K., Kaun, K.R. et Rankin, C.H. (2002). A new group-training procedure for habituation demonstrates that presynaptic glutamate release contributes to long-term memory in *Caenorhabditis elegans*. *Learn. Mem.*, 9(3), 130-137.
- Rose, J.K., Kaun, K.R., Chen, S.H. et Rankin, C.H. (2003). Glutamate receptor trafficking underlies long-term memory in *C. elegans*. *J. Neurosci.*, 23(29), 9595-9600.
- Rosenblatt, F. (1958). The perceptron: a probabilistic model for information storage and organization in the brain. *Psychological Review*, 65, 386-408.
- Rudy, J.W. (2014). *The neurobiology of learning and memory*. (2<sup>e</sup> Ed.) Sinauer Associates, Inc., Sunderland, MA: USA.
- Ruf, B. et Schmidt, M. (1998). Self-organization of spiking neurons using action potential timing. *IEEE Transaction on neural networks*, 9(3), 575-578.
- Rumelhart, D.E., Hinton, G.E. et Williams, R.J. (1986a). Learning representations of back-propagation errors, *Nature (London)*, 323, 533-536.

- Rumelhart, D.E., McClelland, J.L. et The PDP Research Group (1986). *Parallel Distributed Processing: Foundations*, volume 1. Cambridge, MA: MIT Press.
- Rumelhart, D.E., McClelland, J.L. et The PDP Research Group (1986). *Parallel Distributed Processing: Explorations in the microstructure of cognition*, volume 2. Cambridge, MA: MIT Press.
- Russell, S.J. et Norvig, P. (2010). *Artificial Intelligence: A modern approach* (3<sup>e</sup> éd.). Upper Saddle River, NJ, USA, Pearson Education Inc.
- Salotti, J.-M. et Lepretre, F. (2008). Classical and operant conditioning as roots of interaction for robots. Dans: Proceedings of the workshop From motor to interaction learning robots, IEEE International Conference on intelligent robots and systems, Nice.
- Sanz-Clemente, A., Nicoll, R.A. et Roche, K.W. (2013). Diversity in NMDA receptor composition: many regulators, many consequences. *The Neuroscientist*, 19(1), 62-75.
- Schmajuk, N.A. (2010). *Mechanisms in classical conditioning*. Cambridge, UK, Cambridge University Press.
- Schmajuk, N.A. (2010). *Computational models of conditioning*. Cambridge, UK, Cambridge University Press.
- Schultz, W., Dayan, P. et Montague, P.R. (1997). A neural substrate for prediction and reward. *Science*, 275, 1593-1598.
- Sehgal, M., Song, C., Ehlers, V.L., Moyer, J. et James, R. (2013). Learning to learn: Intrinsic plasticity as a metaplasticity mechanism for memory formation. *Neurobiol. Learn. Mem.*, 105, 186-199. doi:10.1016/j.nlm.2013.07.008.
- Sejnowski, T.J. et Rosenberg, C.R. (1987). Parallel networks that learn to pronounce English text. *Complex systems*, 1, 145-168.
- Seth, A.K., Prescott, T.J. et Bryson, J.J. (2011). *Modelling natural action selection*. Cambridge, UK, Cambridge University Press.
- Shouval, H.Z., Bear, M.F. et Cooper, L.N. (2002). A unified theory of NMDA receptor-dependent bidirectional synaptic plasticity. *Proc. Natl. Acad. Sci. USA*, 99, 10831-10836.
- Shouval, H.Z. (2011). What is the appropriate description level for synaptic plasticity? *PNAS*, 108(48), 19103-19104.
- Simcog [s.d.], *AI-Future is a R&D company in the AI field*. Récupéré de <http://aifuture.com>
- Simons-Weidenmaier, N., Weber, M., Pappert, C., Pilz, P.K. et Schmid, S. (2006). Synaptic depression and short-term habituation are located in the sensory part of the mammalian startle pathway. *BMC Neurosci.*, 7:38, doi:10.1186/1471-2202-7-38.
- Simulink [s.d.], *Simulation and Model-Based Design*. Récupéré de <http://mathworks.com/products/simulink>
- Sjöström, P.J., Turrigiano, G.G. et Nelson, S.B. (2001). Rate, timing, and cooperativity jointly determine cortical synaptic plasticity. *Neuron*, 32, 1149-1164.
- Sjöström, P.J. et Nelson, S.B. (2002). Spike timing, calcium signals and synaptic plasticity. *Curr. Opin. Neurobiol.*, 12, 305-314.
- Skinner, B.F. (1938). *The Behavior of Organisms: An Experimental Analysis*. New York: Appleton-Century-Crofts.
- Skorheim, S., Lonjers, P. et Bazhenov, M. (2014). A spiking network model of decision making employing rewarded STDP. *PloSOne*, 9(3), doi: 10.1371/journal.pone.0090821.
- Smythies, J. (2002). *The dynamic neuron: A comprehensive survey of the neurochemical basis of synaptic plasticity*. Cambridge, MA: MIT Press.



- Soltoggio, A. et Steil, J.J. (2013). Solving the distal reward problem with rare correlations. *Neural Computation*, 25(4), 940-978.
- Song, S., Miller, K.D. et Abbott, L.F. (2000). Competitive Hebbian Learning through spike-timing-dependent synaptic plasticity. *Nature Neuroscience*, 3(9), 206-221.
- Sporns, O. (2002). Embodied cognition. Dans: *Handbook of brain theory and neural networks* (2ième édition), Arbib, M. (Dir.), Cambridge MA: MIT Press.
- Staddon, J.E.R. et Higa, J.J. (1996). Multiple time scales in simple habituation. *Psychol. Rev.*, 103(4), 720-733.
- Staddon, J.E.R., Chelaru, I.M. et Higa, J.J. (2002). Habituation, memory and the brain: The dynamics of interval timing. *Behavioural Processes*, 57(2), 71-88.
- Staddon, J.E.R. (2005). Interval timing: memory, not a clock. *TRENDS in Cognitive Sciences*, 9(7), 312-314.
- Stanley, J.C. (1976). Computer simulation of a model of habituation. *Nature*, 261, 146-148.
- Steels, L. et Brooks, R. (1994). *The artificial life route to artificial intelligence: Building Situated Embodied Agents*. New Haven: Lawrence Erlbaum Ass.
- Steidl, S., Rose, J.K. et Rankin, C.H. (2003). Stages of memory in the nematode *Caenorhabditis elegans*. *Behav. Cogn. Neurosci. Rev.*, 2(1), 3-14.
- Steiner, G.Z. et Barry, R.J. (2014). The mechanism of dishabituation. *Frontiers in integrative neuroscience*, 8(14), 1-8.
- Stiles, B.W. et Ghosh, J. (1997). Habituation based neural classifiers for spatio-temporal signals. *Neurocomputing*, 15, 273-307.
- Stopfer, M. et Carew, T.J. (1996). Heterosynaptic facilitation of tail sensory neuron synaptic transmission during habituation in tail-induced and siphon withdrawal reflexes of *Aplysia*. *J. Neurosci.*, 16(16), 4933-4948.
- Sugai, R., Azami, S., Shiga, H., Watanabe, T., Sadamoto, H., Kobayashi, S., Hatakeyama, D., Fujito, Y., Lukowiak, K. et Iro, E. (2007). One-trial conditioned taste aversion in *Lymnaea*: Good and poor performers in long-term memory acquisition. *The journal of experimental Biology*, 210, 1225-1237.
- Suri, R.E. (2002). TD models of reward predictive responses in dopamine neurons. *Neural Networks*, 15, 523-533.
- Suri, R.E. et Schultz, W. (1999). A neural network learns a spatial delayed response task with a dopamine-like reinforcement signal. *Neuroscience*, 91, 871-890.
- Sutton, M.A., Masters, S.E., Bagnall, M.W. et Carew, T.J. (2001). Molecular mechanisms underlying a unique intermediate phase of memory in *Aplysia*. *Neuron*, 31, 143-154.
- Sutton, R.S. (1984). *Temporal credit assignment in reinforcement learning*. PhD thesis, Department of Computer Science, University of Massachusetts, Amherst, MA 01003.
- Sutton, R.S. et Barto, A.G. (1990). Time derivative models of Pavlovian reinforcement. Dans: M. Gabriel et J. Moore (dir.), *Learning and computational neuroscience: Foundations of adaptive networks* (pp. 539-602). Cambridge, MA: MIT Press.
- Sutton, R.S. et Barto, A.G. (1998). *Reinforcement Learning: An Introduction*. MIT Press, Cambridge, MA, USA.
- Szyszkowski, P., Galkin, A. et Menzel, R. (2008). Associative and non-associative plasticity in Kenyon cells of the honeybee mushroom body. *Front. Syst. Neurosci.*, 2, 1-10.
- Thompson, R.F. et Spencer, W.A. (1966). Habituation: a model phenomenon for the study of neuronal substrates of behavior. *Psychol. Rev.*, 73, 16-43.

- Thorndike, E.L. (1901). Animal intelligence: An experimental study of the associative processes in animals. *Psychological Review Monograph Supplement*, 2, 1–109.
- Thorpe, W.H. (1956). *Learning and instinct in animals*. Harvard university press, Cambridge.
- Usrey, W.M., Reppas, J.B. et Reid, R.C. (1998). Paired-spike interactions and synaptic efficacy of retinal inputs to the thalamus. *Nature*, 395, 384–387.
- Valente, A., Huang, K.-H., Portugues, R. et Engert, F. (2012). Ontogeny of classical and operant learning behaviors in zebrafish. *Learn. Mem.*, 19, 170–177.
- Van Gelder, T. (1998). The dynamical hypothesis in cognitive science. *Behavioral and brain sciences*, 21, 615–665.
- Van Leeuwen, M. (2003). *Modelling realistic neurons*. Institute for Information and Computing Sciences, Utrecht University.
- Varela F., Thompson, E. et Rosch, E. (1992). *The Embodied Mind*, Cambridge, MA: MIT Press.
- Vogel, E.H. et Wagner, A.R. (2005). Stimulus specificity in the habituation of the startle response in the rat. *Physiol. Behav.*, 86(4), 516–525.
- Wang, D.L. (1994). Modeling neural mechanisms of vertebrate habituation: Locus specificity and pattern discrimination. *Journal of Computational Neuroscience*, 1(4), 285–299.
- Wang, H.-X., Gerkin, R., Nauen, D. et Wang, G.-Q. (2005). Coactivation and timing-dependent integration of synaptic potentiation and depression. *Nat. Neurosci.*, 8, 187–193.
- Wang, X., Hou, Z.-G., Zou, A., Tan, M. et Cheng, L. (2008). A behavior controller based on spiking neural networks for mobile robots. *Neurocomputing*, 71, 655–666.
- Webb, B. et Consi, T. (2001). *Biorobotics: Methods and applications*. Cambridge, MA, AAAI Press, MIT Press.
- Webots [s.d.], *Webots robot simulator*. Récupéré de <https://www.cyberbotics.com>
- Wicks, S.R. et Rankin, C.H. (1995). Integration of Mechanosensory Stimuli in *Caenorhabditis elegans*. *J. Neurosci.*, 15(3), 2434–2444.
- Wilson, H. R. (1999). *Spikes, decisions and actions: dynamical foundations of neuroscience*. New York: Oxford University Press.
- Ziemke, Tom (2001). Are Robots Embodied? Dans: Balkenius, Zlatev, Brezeal, Dautenhahn & Kozima (Dir.) *Proceedings of the First International Workshop on Epigenetic Robotics: Modeling Cognitive Development in Robotic Systems* (pp. 75–93). Lund University Cognitive Studies, vol. 85, Lund, Sweden.
- Ziemke, T. (2003b). What's that thing called embodiment? Dans: *Proceedings of the 25th Annual Meeting of the Cognitive Science Society*. Boston, MA: Cognitive Science Society.
- Zucker, R.S. et Regehr, W.G. (2002). Short-term synaptic plasticity. *Annu. Rev. Physiol.*, 64, 355–405.